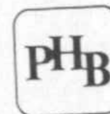
Engenharia de Controle Moderno

Katsuhiko Ogata
UNIVERSIDADE DE MINNESOTA

Tradução de

André Fábio Kohn
José Carlos Teixeira de Barros Moraes
Professores do Departamento de Engenharia de Eletricidade, Escola Politécnica, Universidade de São Paulo



Prentice/Hall do Brasil

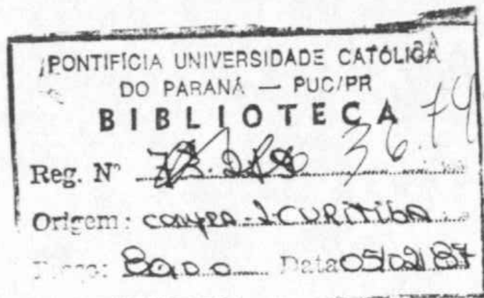
CIP — Brasil. Catalogação-na-fonte
Sindicato Nacional dos Editores de Livros, RJ.

O28e Ogata, Katsuhiko.
Engenharia de controle moderno / Katsuhiko Ogata;tradução de André Fábio Kohn, José Carlos Teixeira de Barros Moraes. — Rio de Janeiro:Prentice-Hall do Brasil, 1985.

Tradução de: Modern control engineering.
Bibliografia.
ISBN 85-7054-019-1

1. Controle automático I. Kohn, André Fábio II. Moraes, José Carlos Teixeira de Barros III. Título

85-0288 CDD — 629.8



Título do original em inglês
Modern Control Engineering

1982 — 5 4 3 2

Fotocomposição da Editora Guanabara Koogan S.A.

Prefácio

Este livro foi preparado para ser um texto completo de um primeiro estudo de engenharia de controle. Está redigido ao nível do último ano de engenharia elétrica, mecânica, aeronáutica ou química.

Este texto abrange tanto a teoria de controle clássico como a de controle moderno. Os três primeiros capítulos apresentam os conceitos fundamentais de sistemas de controle por realimentação e a base matemática elementar necessária para o entendimento do livro. O Cap. 4 trata de modelagem de sistemas físicos. Os próximos seis capítulos (5 a 10) apresentam métodos e técnicas convencionais para analisar e projetar sistemas de controle. Os dois capítulos seguintes (11 e 12) discutem sistemas não lineares. O Cap. 13 dá uma introdução a sistemas de tempo discreto e ao método da transformada z. Os últimos três capítulos (14 a 16) apresentam uma introdução à teoria de controle moderno baseada em conceitos de espaço de estados.

Supõe-se que o leitor esteja familiarizado com equações diferenciais e análise de circuitos em nível introdutório, bem como em mecânica.

Os primeiros 10 capítulos podem ser usados como texto, quer de um curso de um trimestre, quer de um semestre, com três aulas por semana, dependendo de quanta matéria será coberta. Os seis capítulos restantes podem ser usados como texto de um curso de um semestre. No Departamento de Engenharia da Universidade de Minnesota, a matéria deste livro é tratada em dois cursos; a maior parte dos capítulos de 1 a 10 é utilizada em um curso introdutório de um trimestre sobre engenharia de controle, e os Caps. 13 a 16 fazem parte de um curso de um trimestre sobre teoria de controle moderno, ambos os cursos no último ano de engenharia.

Este livro provê muitos exemplos ilustrativos para esclarecer a teoria apresentada. Acredito que a melhor maneira de aprofundar o entendimento da teoria de controle é dominar os fundamentos através da resolução de muitos problemas, cujas soluções estão disponíveis. Partindo deste ponto de vista, além de muitos exemplos, cada capítulo é seguido de inúmeros problemas resolvidos bem como outros não resolvidos. A resolução correta dos problemas não resolvidos demonstrará que o leitor compreendeu a matéria apresentada.

Gostaria de agradecer ao Dr. Richard C. Jordan, chefe do Departamento de Engenharia Mecânica da Universidade de Minnesota, pelo seu contínuo encorajamento, bem como deixar meu reconhecimento a muitos ex-alunos que resolveram muitos exemplos e problemas usados no livro. Gostaria também de expressar meu reconhecimento ao Mr. Richard F. Palas, do Departamento de Engenharia Química da Universidade de Minnesota, por ter feito muitos comentários construtivos bem como melhorado a apresentação do material.

Katsuhiko Ogata



Engenharia de Controle Moderno

Índice

1

Introdução a Sistemas de Controle

1.1 INTRODUÇÃO

Controle automático representa um papel vital no avanço da engenharia e da ciência. Além de possuir importância extrema em sistemas de pilotagem de avião, mísseis guiados, veículos espaciais etc., tornou-se uma parte integrante e importante dos processos industriais e de fabricação modernos. Por exemplo, controle automático é essencial em operações industriais para controle de pressão, temperatura, umidade, viscosidade e fluxo em processos industriais; manuseando, operando e montando partes mecânicas das indústrias de fabricação, entre muitas outras.

Desde que os avanços na teoria e na prática de controle automático propiciam meios para atingir-se desempenho ótimo de sistemas dinâmicos, bem como melhoria na qualidade e diminuição do custo de produção, aumento da taxa de produção, operações manuais repetitivas etc., a maioria dos engenheiros e cientistas deve entender e conhecer bem este campo.

Revisão histórica. O primeiro trabalho significativo em controle automático foi o de James Watt, que construiu um controlador centrífugo para controle de velocidade de uma máquina a vapor no século XVIII. Outros trabalhos importantes nos primeiros estágios de desenvolvimento da teoria de controle são os de Minorsky, Hazen e Nyquist, entre muitos outros. Em 1922, Minorsky trabalhou em controladores automáticos para pilotagem de navios e mostrou como poderia ser "determinada" a estabilidade a partir das equações diferenciais que descrevem o sistema. Em 1932, Nyquist desenvolveu um procedimento relativamente simples para determinar a estabilidade de sistemas de malha-fechada com base na resposta a entradas senoidais em regime permanente da malha-aberta. Em 1934, Hazen, que introduziu o termo "servomecanismos" para sistemas de controle de posição, discutiu o projeto de servomecanismos a relé capazes de seguir muito de perto uma entrada variável.

Durante a década de 1940, os métodos de resposta em freqüência tornaram possível aos engenheiros projetar sistemas de controle realimentados lineares, que satisfaziam os requisitos de desempenho. Desde o final da década de 1940 até o

início dos anos 50, o método do lugar das raízes em projeto de sistemas de controle foi completamente desenvolvido.

Os métodos de resposta em freqüência e lugar das raízes que correspondem ao coração da teoria de controle clássica levaram a sistemas que são estáveis e satisfazem um conjunto de requisitos de desempenho mais ou menos arbitrários. Estes sistemas não são, em geral, ótimos no sentido lato. Desde a década de 1950, a ênfase nos problemas de projeto de controle tem sido transferida do projeto de um dos muitos sistemas que operam para o projeto de um sistema ótimo em algum sentido lato.

Em virtude de os processos modernos com muitas entradas e saídas tornarem-se mais e mais complexos, a descrição de um sistema de controle moderno exige um grande número de equações. A teoria de controle clássica, que trata apenas de sistemas de entrada-simples-saída-simples, tornou-se inteiramente impotente para sistemas de múltiplas-entradas-múltiplas-saídas. Desde 1960, aproximadamente, a teoria de controle moderna tem sido desenvolvida para competir com a complexidade crescente de processos modernos e requisitos rigorosos e estreitos em precisão, peso e custo em aplicações militares, espaciais e industriais.

Devido à real disponibilidade de computadores digitais, analógicos e híbridos eletrônicos para uso em cálculos complexos, a utilização de computadores no projeto de sistemas de controle e o uso de computadores *on-line* na operação de sistemas de controle constituem atualmente uma prática comum.

Pode-se dizer que os desenvolvimentos mais recentes na teoria de controle moderna estão na direção do controle ótimo tanto de sistemas determinísticos como estocásticos, bem como para controle de aprendizado e adaptativo de sistemas complexos. Aplicações da teoria de controle moderna em áreas não de engenharia, tais como biologia, economia, medicina e sociologia, estão em desenvolvimento, e resultados interessantes e significativos podem ser esperados em um futuro próximo.

1.2 DEFINIÇÕES

Nesta seção definiremos a terminologia necessária para descrever sistemas de controle.

Plantas (*Plants*)*. Uma planta é uma parte de equipamento, eventualmente um conjunto de itens de uma máquina, que funcionam conjuntamente, cuja finalidade é desempenhar uma dada operação. Neste livro designaremos qualquer objeto físico a ser controlado (tais como uma nave espacial, um reator químico ou uma caldeira para aquecimento) como uma *planta*.

Processos (*Processes*)*. O Dicionário Merriam-Webster define um processo como uma operação ou desenvolvimento natural, que evolui progressivamente, caracterizado por uma série de mudanças graduais que se sucedem, uma em relação às outras, de um modo relativamente fixo e objetivando um particular resultado ou meta; ou uma operação artificial ou voluntária, que evolui progressivamente e se constitui por uma série de ações controladas ou movimentos sistematicamente dirigidos objetivando um particular resultado ou meta. Neste livro designaremos qualquer operação a ser controlada como um *processo*. Exemplos são processos químicos, econômicos e biológicos.

*N. do T.: Devido à inexistência na língua portuguesa do vocábulo "planta" com a conotação necessária em Sistema de Controle, utilizar-se-á preferencialmente o termo "processo" englobando *plant* e *process*, a não ser onde possa causar dificuldade de interpretação.

Sistemas. Um sistema é uma combinação de componentes que atuam conjuntamente e realizam um certo objetivo. Um sistema não é limitado a algo físico. O conceito de sistema pode ser aplicado a fenômenos abstratos, dinâmicos, tais como os encontrados em economia. A palavra "sistema" deve, portanto, ser interpretada para designar sistemas físicos, biológicos, econômicos etc.

Distúrbios. Um distúrbio é um sinal que tende a afetar adversamente o valor da saída de um sistema. Se um distúrbio é gerado dentro do sistema, ele é denominado *interno*; ao passo que um distúrbio externo é gerado fora do sistema e constitui uma entrada.

Controle realimentado. Controle realimentado é uma operação que, na presença de distúrbios, tende a reduzir a diferença entre a saída de um sistema e a entrada de referência (ou um estado desejado, arbitrariamente variado) e que opera com base nesta diferença. Aqui, apenas distúrbios não previsíveis (isto é, aqueles não conhecidos *a priori*) são designados como tais, desde que com distúrbios conhecidos ou previsíveis é sempre possível incluir compensação dentro do sistema de modo que aquelas medidas sejam desnecessárias.

Sistemas de controle realimentados. Um sistema de controle realimentado é aquele que tende a manter uma relação prescrita entre a saída e a entrada de referência, comparando-as e utilizando a diferença como um meio de controle.

Note que os sistemas de controle realimentados não são limitados ao campo da engenharia mas podem ser encontrados em várias outras áreas, tais como economia e biologia. Por exemplo, o organismo humano, sob certo aspecto, é análogo a uma planta química intrincada com uma grande variedade de operações unitárias. O controle de processo desta rede de reação química e transporte envolve uma variedade de laços de controle. De fato, o organismo humano é um sistema de controle realimentado extremamente complexo.

Servomecanismos. Um servomecanismo é um sistema de controle realimentado no qual a saída é alguma posição mecânica, velocidade ou aceleração. Portanto, os termos *servomecanismos* e *sistema de controle de posição* (ou *de velocidade*, ou *de aceleração*) são sinônimos. Servomecanismos são muito usados na indústria moderna. Por exemplo, a operação automática completa de máquinas operatrizes, juntamente com instruções programadas, pode ser desempenhada com o uso de servomecanismos.

Sistemas reguladores automáticos. Um sistema regulador automático é um sistema de controle realimentado no qual a entrada de referência, ou a saída desejada, ou é constante ou varia lentamente com o tempo, e no qual a finalidade principal é manter a saída real em um valor desejado, na presença de distúrbios.

Um sistema de aquecimento residencial que utiliza um termostato como controlador é um exemplo de um sistema regulador automático. Neste sistema, o ajuste do termostato (a temperatura desejada) é comparado com a temperatura real do ambiente. Uma variação na temperatura externa é um distúrbio neste sistema. O objetivo é manter a temperatura desejada no ambiente fixa e independente de variações na temperatura externa. Há muitos outros exemplos de sistemas reguladores automáticos, alguns dos quais são os controles automáticos de pressão e de grandezas elétricas tais como tensão, corrente e freqüência.

Sistemas de controles de processos. Um sistema regulador automático no qual a saída é uma variável tal como uma temperatura, pressão, fluxo, nível de líquido ou pH é denominado um *sistema de controle de processo*. Controle de processo é exaustivamente aplicado na indústria. Controles programados, tal como controle de temperatura de fornos para aquecimento nos quais a temperatura do forno é controlada de acordo com um programa pré-selecionado, são muitas vezes usados nestes sistemas. Por exemplo, um programa pré-selecionado pode ser tal que a temperatura do forno (ou, eventualmente, em uma caldeira) é aumentada até uma dada temperatura, em um dado intervalo de tempo, e então diminui para uma outra dada temperatura em algum outro intervalo de tempo conhecido. Em um controle programado como este, o ponto de ajuste pode ser variado de acordo com o tempo preestabelecido desejado. O controlador, então, tem a função de manter a temperatura no forno próxima ao ponto de ajuste variável. Deve ser observado que muitos sistemas de controle de processos incluem servomecanismos como uma parte integral.

1.3 CONTROLE EM MALHA-FECHADA E CONTROLE EM MALHA-ABERTA

Definiremos inicialmente sistemas de controle em malha-fechada e em malha-aberta. Faremos posteriormente uma comparação destes dois tipos. Finalmente, serão introduzidos os conceitos de controle adaptativo e controle de aprendizado.

Sistema de controle em malha-fechada. Um sistema de controle em malha-fechada é aquele no qual o sinal de saída possui um efeito direto na ação de controle. Isto é, sistemas de controle em malha-fechada são sistemas de controle realimentados. O sinal erro atuante, que é a diferença entre o sinal de entrada e o sinal realimentado (que pode ser o sinal de saída ou uma função do sinal de saída e suas derivadas), é alimentado no controlador de modo a reduzir o erro e manter a saída do sistema em um valor desejado. Em outras palavras, o termo "malha-fechada" implica o uso de ação de realimentação com a finalidade de reduzir o erro do sistema. A Fig. 1.1 mostra a relação entrada-saída do sistema de controle em malha-fechada. A representação indicada na figura é denominada *diagrama de blocos*. Para ilustrar o conceito de sistemas de controle em malha-fechada, considere o sistema térmico mostrado na Fig. 1.2, onde um ser humano atua como o controlador. Ele deseja manter a temperatura da água quente em um dado valor. O termômetro instalado na tubulação de saída da água quente mede a temperatura real. Esta temperatura é a saída do sistema. Se o operador observa o termômetro e

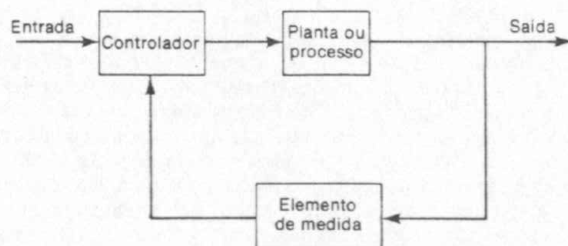


Fig. 1.1 Sistema de controle em malha-fechada.

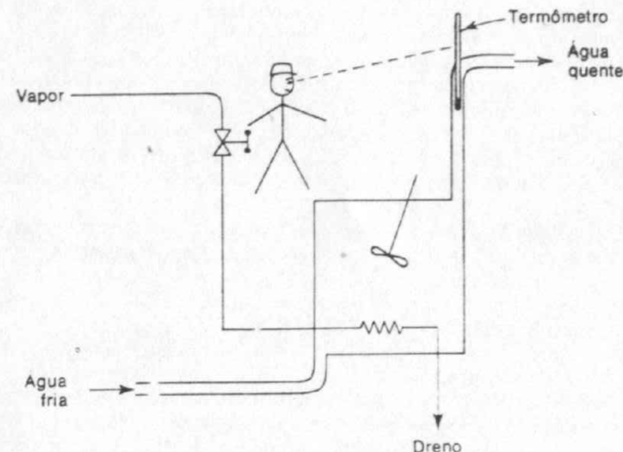


Fig. 1.2 Controle realimentado manual de um sistema térmico.

verifica que a temperatura é maior do que a desejada, ele reduz a quantidade de suprimento de vapor de modo a diminuir esta temperatura. É bem possível que a temperatura se torne demasiado baixa, necessitando repetir a seqüência de operações no sentido oposto.

Esta ação de controle é baseada na operação de malha-fechada. Desde que tanto a realimentação da saída (temperatura da água), para comparação com a entrada de referência, como a ação de controle ocorrem através de ações do operador, este é um sistema de controle em malha-fechada. Sistemas como este podem ser denominados sistemas de controle em malha-fechada manuais ou com realimentação manual.

Se for utilizado um controlador automático para substituir o operador humano, conforme é mostrado na Fig. 1.3, o sistema de controle torna-se automático, isto é, um sistema de controle em malha-fechada automático ou com realimentação automática. A posição do dispositivo de seleção de temperatura (um *dial*, por exemplo) no controlador automático seleciona a temperatura desejada. A saída, a temperatura real da água quente, que é medida pelo dispositivo de medida de temperatura, é comparada com a temperatura desejada de modo a gerar um sinal de erro atuante. Desta forma, a temperatura de saída é convertida nas mesmas unidades da entrada (ponto de ajuste) por meio de um transdutor. (Um transdutor é um dispositivo que converte um sinal de uma forma para outra.) O sinal erro produzido no controlador automático é amplificado, e a saída do controlador é enviada a uma válvula de controle para variar a abertura da válvula e, conseqüentemente, o suprimento de vapor de modo a corrigir a temperatura real da água. Se não houver erro, nenhuma variação é necessária na abertura da válvula.

Nos sistemas aqui considerados, as variações na temperatura ambiente, a temperatura de água fria na tubulação de entrada, etc. podem ser considerados distúrbios externos.

Os sistemas de controle com realimentação manual e realimentação automática anteriormente citados operam de maneira similar. Os olhos do operador

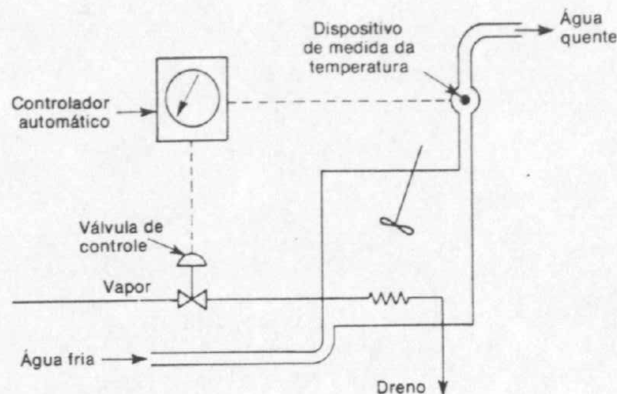


Fig. 1.3 Controle realimentado automático de um sistema térmi :o.

constituem o dispositivo análogo ao medidor de erro; seu cérebro, o análogo do controlador automático; e seus músculos, o análogo do atuador.

O controle de um sistema complexo por um operador humano não é eficiente devido às inúmeras inter-relações entre as diversas variáveis. Note que mesmo em um sistema simples, um controlador automático eliminará quaisquer erros humanos de operação. Se for necessária alta precisão de controle, o controle deve ser automático.

Inúmeros sistemas de controle em malha-fechada podem ser encontrados na indústria e em residências. Alguns exemplos são todos os servomecanismos, a maioria dos sistemas de controle de processos, refrigeradores residenciais, aquecedores de água automáticos e sistemas de aquecimento residenciais automáticos com controle termostático.

Sistemas de controle em malha-aberta. Sistemas de controle em malha-aberta são sistemas de controle nos quais a saída não tem efeito na ação do controle. Isto é, em um sistema de controle em malha-aberta, a saída nem é medida nem é realimentada para comparação com a entrada. A Fig. 1.4 indica a relação entrada-saída de um sistema deste tipo. Um exemplo prático é uma máquina de lavar roupa. As operações de molhar, lavar e enxaguar em uma máquina de lavar roupa são efetuadas em uma mesma base de tempo. A máquina não mede o sinal de saída, isto é, a limpeza das roupas.

Em qualquer sistema de controle em malha-aberta a saída não é comparada com a entrada de referência. Conseqüentemente, a cada entrada de referência corresponde uma condição de operação fixa. Ou seja, a precisão do sistema depende de uma calibração. (Sistemas de controle em malha-aberta devem ser cuidadosamente calibrados e devem manter esta calibração de forma a serem úteis.) Na presença de distúrbios, um sistema de controle em malha-aberta não desempenhará a tarefa desejada. O controle de malha-aberta, na prática, somente pode ser usado se a relação entre a entrada e a saída for conhecida e não houver distúrbios internos ou externos. É claro que estes sistemas não são sistemas de controle realimentados. Note que qualquer sistema de controle que opera em uma base de tempo é de malha-aberta. Por exemplo, controle de tráfego por meio de sinais operados em uma base de tempo é um outro tipo de controle em malha-aberta.

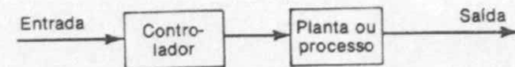


Fig. 1.4 Sistema de controle em malha-aberta.

Sistemas de controle em malha-fechada *versus* malha-aberta. Uma vantagem do sistema de controle em malha-fechada é que o uso de realimentação torna a resposta do sistema relativamente insensível a distúrbios externos e variações internas em parâmetros do sistema. É portanto possível a utilização de componentes baratos e sem muita precisão para obter o controle preciso de um dado processo; isto é impossível no caso de malha-aberta.

Do ponto de vista de estabilidade, é mais fácil construir o sistema de controle em malha-aberta desde que a estabilidade não constitui um problema significativo. Por outro lado, a estabilidade é sempre um problema fundamental no sistema de controle em malha-fechada já que pode tender a corrigir erros e causar oscilações de amplitudes constante ou variável.

Deve ser enfatizado que, para sistemas nos quais as entradas são conhecidas antecipadamente no tempo e nos quais não há distúrbios, é aconselhável usar controle em malha-aberta. Sistemas de controle em malha-fechada possuem vantagens apenas quando distúrbios imprevisíveis e/ou variações imprevisíveis em componentes do sistema estão presentes. Note que a potência de saída determina parcialmente o custo, peso e dimensão de um servomecanismo (ou investimento de capital, capacidade humana etc. em um sistema comercial). De maneira a diminuir a potência necessária de um sistema, pode ser utilizado controle em malha-aberta, quando aplicável. Uma combinação apropriada de controle em malha-aberta e em malha-fechada normalmente é mais barata e fornece um desempenho global do sistema bastante satisfatório.

Controles direto *versus* indireto. Note que para obter o melhor resultado é desejável medir e controlar diretamente as variáveis que indicam o estado do sistema ou a qualidade do produto. No caso de sistemas de controle de processos podemos querer medir e controlar diretamente a qualidade do produto. Entretanto, isto pode constituir um problema difícil, desde que esta qualidade pode ser difícil de medir. Se este for o caso, torna-se necessário controlar uma variável secundária. Por exemplo, variáveis (tais como temperatura e pressão) que são diretamente relacionadas com a qualidade podem ser controladas. Como outras variáveis podem afetar a relação entre a qualidade e a variável medida, o controle indireto de um sistema não é normalmente tão eficiente como o controle direto. Embora possa ser difícil, sempre deve ser tentado controlar a variável primária tão diretamente quanto possível.

Sistemas de controle adaptativos. As características dinâmicas da maioria dos sistemas de controle não são constantes devido a várias razões, tais como deterioração de componentes ao longo do tempo ou variações em parâmetros e ambiente (por exemplo, variações na massa e condições atmosféricas de um sistema de controle de nave espacial). Embora os efeitos de pequenas variações sobre as características dinâmicas sejam atenuados em um sistema de controle realimentado, se as variações nos parâmetros do sistema e ambiente forem significativas, um sistema satisfatório deve possuir a habilidade de adaptação. A adaptação implica a habilidade de se auto-ajustar ou automodificar de acordo com variações imprevisíveis nas condições de ambiente ou estrutura. Os sistemas de controle que possuem

uma habilidade franca de adaptação são denominados sistemas de controle *adaptativos*.

Em um sistema de controle adaptativo, as características dinâmicas devem ser identificadas em todos os instantes de modo que os parâmetros do controlador possam ser ajustados de maneira a manter desempenho ótimo. Este conceito possui um grande atrativo para o projetista de sistema desde que um sistema de controle adaptativo, além de acomodar variações ambientais, também acomodará erros ou incertezas de projeto de engenharia moderados e compensará falhas de componentes do sistema de pequena monta, aumentando conseqüentemente a confiabilidade global do sistema.

Sistemas de controle de aprendizado. Muitos sistemas de controle, aparentemente de malha-aberta, podem ser convertidos em um sistema de controle em malha-fechada se um operador humano é considerado como controlador, comparando a entrada e a saída e efetuando a ação corretiva baseada no erro ou diferença resultante.

Se tentarmos analisar estes sistemas de controle em malha-fechada operados-pelo-homem, encontraremos o difícil problema de escrever as equações que descrevem o comportamento de um ser humano. Um dos muitos fatores complicados neste caso é a habilidade de aprendizado do operador humano. Conforme o operador ganha mais experiência, ele se tornará melhor controlador, e isto deve ser levado em consideração na análise de tal sistema. Sistemas de controle que possuem uma habilidade para aprender são denominados sistemas de controle *de aprendizado*. Este conceito é relativamente novo e ainda não foi completamente explorado.

Recentes progressos em aplicações de controle adaptativos e de aprendizado foram relatados na literatura, porém grandes segmentos da atividade de engenharia permanecem disponíveis para futuros estudos.

1.4 EXEMPLOS ILUSTRATIVOS DE SISTEMAS DE CONTROLE

Nesta seção apresentaremos alguns exemplos ilustrativos de sistemas de controle em malha-fechada.

Sistemas de controle de pressão. A Fig. 1.5 mostra um sistema de controle de pressão. A pressão na caldeira é controlada pela posição do regulador. Esta pressão é medida por um elemento medidor de pressão. O sinal então obtido é enviado ao controlador para comparação com o valor desejado. Se houver qualquer diferença ou erro, a saída do controlador é enviada ao atuador que posiciona o regulador convenientemente de modo a reduzir o erro.

Sistemas de controle de velocidade. O princípio básico do regulador de Watt para máquinas a vapor é ilustrado no diagrama esquemático da Fig. 1.6. A quantidade de vapor admitida no cilindro da máquina é ajustada de acordo com a diferença entre as velocidades real e desejada.

A seqüência de ações pode ser estabelecida como segue: a entrada de referência (ponto de ajuste) é selecionada de acordo com a velocidade desejada. Se a velocidade real cair abaixo da velocidade desejada, então o decréscimo na força centrífuga do regulador de velocidade faz com que a válvula de controle se mova para cima, suprindo mais vapor, e a velocidade da máquina aumenta até o valor desejado ser atingido. Por outro lado, se a velocidade da máquina aumentar acima da velocidade desejada, o aumento na força centrífuga do regulador causa um movimento para baixo da válvula de controle. Este movimento diminui o suprimento de vapor e a velocidade da máquina se reduz até atingir o valor desejado.

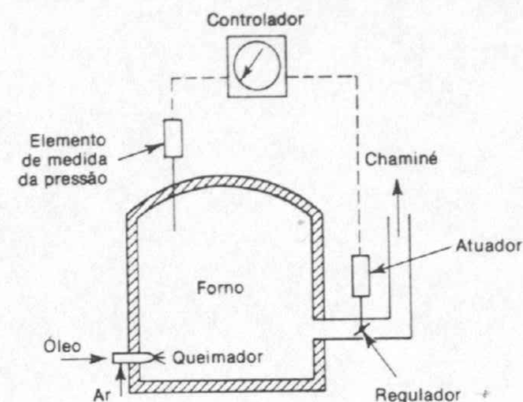


Fig. 1.5 Sistema de controle de pressão.

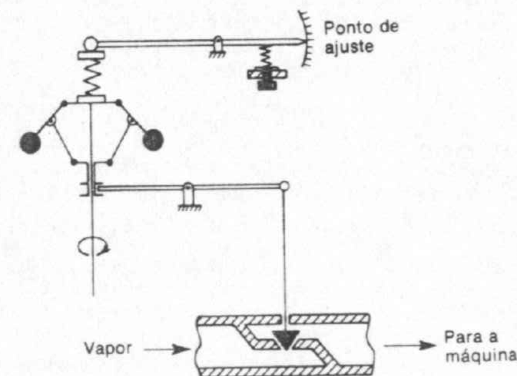


Fig. 1.6 Sistema de controle de velocidade.

Sistemas de controle numéricos. Controle numérico é um método de controle dos movimentos de componentes de máquinas pelo uso de números. Em controle numérico o movimento de uma ferramenta pode ser controlado pela informação binária contida em uma fita.

Neste controle, valores numéricos simbólicos são convertidos em valores físicos (dimensões ou grandezas) por meio de sinais elétricos (ou outros) que são deslocados em um movimento linear ou circular. Estes sinais ou são digitais (pulsos) ou analógicos (tensões variáveis no tempo).

O sistema mostrado na Fig. 1.7 opera como segue: uma fita é preparada na forma binária representando a parte P desejada. Para iniciar o sistema, a fita é introduzida na leitora. O sinal de pulso de entrada modulado em freqüência é comparado com o sinal de pulso realimentado. O conversor digital-analógico converte o pulso em um sinal analógico que representa um certo valor de tensão a qual

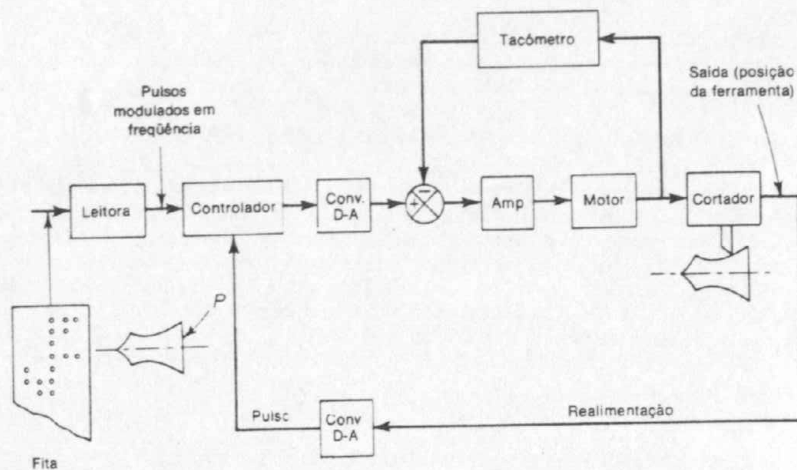


Fig. 1.7 Controle numérico de uma máquina.

por sua vez acarreta a rotação do servomotor. A posição da ferramenta é controlada de acordo com a entrada do servomotor. O transdutor acoplado à ferramenta converte o movimento em um sinal elétrico que é convertido a um sinal pulsado por meio de um conversor analógico-digital. Então este sinal é comparado com o sinal de pulso de entrada. O controlador elimina operações matemáticas nas diferenças dos sinais pulsados. Se houver qualquer diferença entre estes dois, um sinal é enviado ao servomotor para reduzi-lo.

Uma vantagem do controle numérico é que partes complexas podem ser produzidas com tolerâncias uniformes na máxima velocidade de perfuração.

Sistemas de controle por computador. A Fig. 1.8 mostra um diagrama esquemático do controle por computador de um autoforno. O autoforno é uma estrutura enorme com aproximadamente 30 metros de altura. Fornos modernos são construídos para produzir mais de 4.000 toneladas de lingote de ferro por dia e devem ser considerados em operação contínua devido à natureza do processo de fusão.

O minério de ferro, coque e pedra calcária são carregados no forno em proporções apropriadas. (Aproximadamente 2 toneladas de minério, 1 tonelada de coque, $^{1}/_{2}$ tonelada de fluxo e $4^{1}/_{2}$ toneladas de ar são necessários para produzir 1 tonelada de lingote de ferro.) O ar, muito importante neste processo, é aquecido em fornos e injetado na fornalha. O calor na fornalha é produzido pela combustão do coque, da qual o gás de monóxido de carbono é gerado pelo processo de combustão parcial. Este gás, junto com o coque, reduz o minério de ferro na fornalha a um metal, e a pedra calcária atua como um fluxo, eliminando as impurezas. O ferro fundido escoa para a base da fornalha e as impurezas líquidas sobem para a superfície. O ferro fundido e as impurezas líquidas são periodicamente drenados para fora através de orifícios construídos para esta finalidade.

Desde que a presença de carbono, manganês, silício, enxofre, fósforo etc., depende muito da composição do minério, coque, e pedra calcária usados, é muito difícil para operadores humanos controlarem a composição química do lingote de ferro proveniente da fornalha.

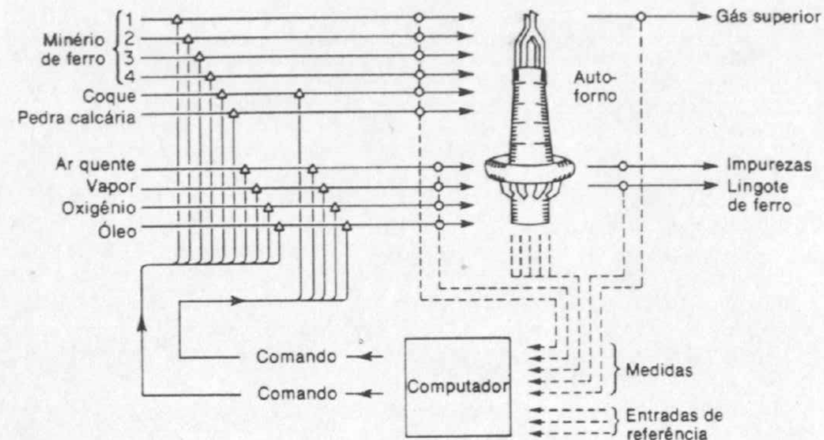


Fig. 1.8 Controle por computador de um autoforno.

No controle por computador destes fornos, as informações sobre composições de lingote de ferro, impurezas, gás, temperatura e pressão no forno, bem como composições do minério, coque e pedra calcária, são alimentadas no computador em certos intervalos de tempo. Os cálculos complexos para a determinação das quantidades ótimas dos vários materiais envolvidos a fim de se carregar o forno são efetuados pelo computador. Então é possível manter a composição do lingote de ferro conforme desejado. Também é possível manter a operação em regime permanente do autoforno em uma condição satisfatória.

Note que em um controle por computador deste processo é necessário possuir modelos matemáticos disponíveis. Porém a dedução de modelos convenientes pode ser um problema difícil pois não são conhecidos todos os fatores que afetam a dinâmica do sistema. Note também que a medida de todas as variáveis necessárias para o controle por computador pode ser difícil ou impossível, e nos casos das variáveis não mensuráveis devem-se fazer estimativas através de métodos estatísticos.

Sistemas de controle de tráfego. Conforme foi estabelecido na Seção 1.3, o controle de tráfego por meio de sinais de tráfego operados em uma base de tempo constitui um sistema de controle de malha-aberta. Se, entretanto, o número de carros parados em cada sinal de tráfego de uma área congestionada em uma cidade for medido continuamente e a informação enviada a um computador de controle central que controla os sinais de tráfego, então o sistema passa a ser de malha-fechada.

O movimento de tráfego em redes é muito complexo devido à variação no volume de tráfego depender fortemente da hora e do dia da semana, bem como de muitos outros fatores. Em alguns casos a distribuição de Poisson pode ser admitida para as chegadas de carros nos cruzamentos, porém isto não é necessariamente válido para todos os problemas de tráfego. De fato, a minimização do tempo de espera médio é um problema de controle muito complexo.

Sistemas biológicos. Considere a competição de duas espécies de bactérias cujas populações são x_1 e x_2. As duas estão competindo no sentido em que consomem a mesma fonte de alimento. Sob certas condições, as populações x_1 e x_2 variam com o tempo de acordo com

$$\dot{x}_1 = a_{11}x_1 - a_{12}x_1x_2$$
$$\dot{x}_2 = a_{21}x_2 - a_{22}x_1x_2$$

onde a_{11}, a_{12}, a_{21} e a_{22} são constantes positivas e x_1 e x_2 são não negativas. Estas equações são as chamadas equações de competição de Volterra.

Se um dado agente químico é fornecido às espécies, as populações variam de acordo com as seguintes equações:

$$\dot{x}_1 = a_{11}x_1 - a_{12}x_1x_2 - b_1u$$
$$\dot{x}_2 = a_{21}x_2 - a_{22}x_1x_2 - b_2u$$

onde b_1 e b_2 são constantes positivas e u é a entrada controlante (a quantidade de agente químico neste exemplo). Um problema interessante é minimizar, em um dado intervalo de tempo, a população x_1 enquanto se mantém a população x_2 tão grande quanto possível. Este é um exemplo de um sistema biológico no qual pode ser aplicada a teoria de controle.

Sistemas de controle de estoque. A programação industrial da taxa de produção e nível de estoque é um outro exemplo de sistema de controle em malha-fechada. O nível de estoque real, que é a saída do sistema, é comparado com o nível de estoque desejado, que pode variar de tempos em tempos de acordo com o mercado. Se houver qualquer diferença entre o nível de estoque real e o nível de estoque desejado, então a taxa de produção é ajustada de modo que a saída sempre esteja no nível desejado ou próximo dele, o qual é escolhido para maximizar o lucro.

Sistemas comerciais. Um sistema comercial pode consistir em muitos grupos. Cada tarefa designada a um grupo representará um elemento dinâmico do sistema. Métodos com realimentação para relato das realizações de cada grupo devem ser estabelecidos neste sistema a fim de se obter uma operação apropriada. O inter-relacionamento entre grupos funcionais deve ser minimizado de modo a reduzir os atrasos indesejáveis no sistema. Quanto menor este inter-relacionamento, menores serão o fluxo do sinais e materiais de trabalho.

Um sistema comercial é um sistema de malha-fechada. Um bom projeto reduzirá o controle administrativo requerido. Note que os distúrbios neste sistema correspondem à carência de mão-de-obra ou materiais, interrupção de comunicação, erros humanos etc.

O estabelecimento de um sistema de estimação bem fundamentado baseado em estatística é imperativo para uma administração apropriada. (Note que é um fato bem conhecido que o desempenho de um sistema deste tipo pode ser melhorado pelo uso de "tempo adiantado" ou "antecipação".)

Para aplicar a teoria de controle a fim de melhorar o desempenho de um sistema deste tipo, devemos representar as características dinâmicas dos grupos componentes do sistema por um conjunto de equações relativamente simples.

Embora seja certamente um problema difícil deduzir as representações matemáticas dos grupos componentes, a aplicação de técnicas de otimização a sistemas comerciais representa uma área interessante na qual melhorias significativas em desempenho podem ser esperadas.

1.5 PRINCÍPIOS DE PROJETO EM SISTEMAS DE CONTROLE

Requisitos gerais de um sistema de controle. Qualquer sistema de controle deve ser estável. Esta é uma exigência fundamental. Além da estabilidade absoluta, um sistema de controle deve possuir uma estabilidade relativa razoável; isto é, a velocidade de resposta deve ser razoavelmente rápida e esta resposta deve possuir uma amortecimento razoável. Um sistema de controle também deve ser capaz de reduzir erros a zero ou a algum valor pequeno tolerável. Qualquer sistema de controle útil deve satisfazer estes requisitos.

A exigência de estabilidade relativa razoável e precisão em estado estacionário tendem a ser incompatíveis. Nos projetos de sistemas de controle deve-se, portanto, estabelecer o compromisso mais eficiente entre estes dois requisitos.

Problemas básicos no projeto de sistema de controle. A Fig. 1.9 representa um diagrama de blocos de um sistema de controle. O controlador produz sinais de controle baseado nas variáveis de entrada de referência e nas variáveis de saída. Nas situações práticas sempre haverá alguns distúrbios agindo na planta. Estes podem ser de origem externa ou interna e aleatórios ou previsíveis. O controlador deve levar em consideração quaisquer distúrbios que possam afetar as variáveis de saída.

Para determinar o sinal de controle ótimo é necessário definir o índice de desempenho. Este índice é uma medida quantitativa do desempenho, medindo o desvio em relação ao desempenho ideal. A especificação do sinal de controle durante o intervalo de tempo de operação é denominada a *lei de controle*. Matematicamente, o problema de controle básico é determinar a lei de controle ótimo, sujeita a vários vínculos de engenharia e econômicos, que minimizam um dado índice de desempenho.

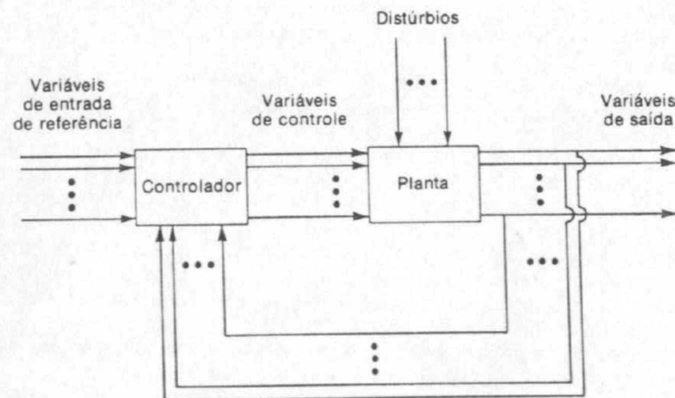


Fig. 1.9 Diagrama de blocos de um sistema de controle.

Para sistemas relativamente simples, a lei de controle pode ser obtida analiticamente. Para sistemas complexos pode ser necessário gerar uma lei de controle ótimo com um computador digital *on-line*.

Análise. Por análise de um sistema de controle designamos a investigação, sob condições específicas do desempenho, do sistema cujo modelo matemático é conhecido.

Desde que qualquer sistema é constituído por componentes, a análise deve iniciar com uma descrição matemática de cada componente. Uma vez que o modelo matemático do sistema completo foi obtido, a maneira pela qual a análise é conduzida independe de o sistema físico ser pneumático, elétrico, mecânico etc.

Projeto. Projetar um sistema significa determiná-lo de modo a desempenhar uma dada tarefa. Em geral o procedimento de projeto não é direto e requer métodos de tentativa-e-erro.

Síntese. Por síntese, interpretamos a determinação por um procedimento direto de um sistema que funcione de uma maneira específica. Normalmente este procedimento é inteiramente matemático desde seu início até o fim do processo do projeto. Procedimentos de síntese são atualmente disponíveis para redes lineares e para sistemas lineares ótimos.

Abordagem básica em projetos de sistemas de controle. A abordagem básica para o projeto de qualquer sistema de controle prático envolve, necessariamente, procedimentos de tentativa-e-erro. A síntese de sistemas de controle lineares é teoricamente possível, e o engenheiro de controle pode determinar sistematicamente os componentes necessários para desempenhar uma dada função e atingir o objetivo desejado. Na prática, entretanto, o sistema pode estar sujeito a muitos vínculos ou pode ser não linear, e, nestes casos, não existe nenhum método de síntese disponível até o presente. Além disso, as características dos componentes podem não ser conhecidas com precisão. Conseqüentemente, procedimentos de tentativa-e-erro sempre são necessários.

Na prática quase sempre são encontradas situações onde é fornecida uma certa planta, e o engenheiro de controle deve projetar o resto do sistema de modo que o conjunto apresente as especificações desejadas desempenhando uma dada tarefa. Note que as especificações devem ser interpretadas em termos matemáticos.

É importante lembrar que algumas das especificações podem não ser realistas. Neste caso, as especificações devem ser revistas nos estágios iniciais do projeto.

Em muitos casos, o projeto de um sistema de controle se desenvolve como segue: o engenheiro inicia o procedimento do projeto conhecendo as especificações ou índice de desempenho, a dinâmica da planta fornecida e a dinâmica dos componentes; este último envolve parâmetros de projeto. O engenheiro de controle aplica então, técnicas de síntese, se disponíveis, conjuntamente com outras técnicas a fim de elaborar um modelo matemático do sistema.

Uma vez formulado o problema do projeto em termos deste modelo, o engenheiro desenvolve um projeto matemático que fornece a solução da versão matemática do problema do projeto. (Neste estágio, é importante a simulação do modelo matemático em um computador. Note que a teoria de controle ótimo é muito útil neste estágio de projeto porque fornece o limite superior do desempenho do sistema para um dado índice de desempenho.)

Após haver sido completado o projeto matemático, o engenheiro de controle

simula um modelo em um computador a fim de testar o comportamento do sistema resultante em resposta a vários sinais e distúrbios. Normalmente, a configuração do sistema inicial não é satisfatória. O sistema deve então ser reprojetado e efetuada a análise correspondente. Este procedimento de projeto e análise é repetido até ser obtido um sistema satisfatório. Pode então ser construído um sistema físico protótipo.

Note que este procedimento para construir um protótipo é o inverso daquele utilizado na modelagem. O protótipo é um sistema físico que representa o modelo matemático com precisão razoável. Uma vez que o protótipo foi construído, o engenheiro o testa a fim de verificar se é ou não satisfatório. Se for, o projeto está concluído. Em caso contrário, o protótipo deve ser modificado e testado. Este procedimento continua até o protótipo estar completamente satisfatório.

1.6 ESBOÇO DO TEXTO

No projeto de um sistema de controle, o engenheiro de controle deve estar apto a determinar e analisar a resposta do sistema em relação a vários sinais e distúrbios. Portanto, iniciaremos com estes problemas de análise. Os Caps. 2-9 e 11-15 são fundamentalmente relacionados com estes problemas, ao passo que os Caps. 10 e 16 dizem respeito exclusivamente a problemas de projeto. Para auxiliar o leitor, discutiremos brevemente o conteúdo e a organização deste livro.

Os Caps. 2 e 3 apresentam uma revisão da base matemática necessária para o entendimento da teoria de controle apresentada neste livro.

O Cap. 4 discute os modelos matemáticos de sistemas físicos. Inicialmente definimos função de transferência e diagrama de bloco, apresentamos uma revisão de várias leis físicas básicas, e deduzimos as equações diferenciais e funções de transferência para vários sistemas físicos. Este capítulo inclui uma discussão básica de gráficos de fluxo de sinal.

O Cap. 5 apresenta vários modos de controle e a introdução de mecânica dos fluidos. O Cap. 6 é relacionado com a análise no domínio do tempo de sistemas de controle. É investigada em detalhes a resposta transitória de sistemas de controle. O critério de estabilidade de Routh é incluído neste capítulo. Também é incluída uma breve discussão sobre técnicas de simulação com computador analógico.

O Cap. 7 apresenta análises de erros e uma introdução à otimização de sistemas. O Cap. 8 introduz o método do lugar das raízes, enquanto o Cap. 9 apresenta as técnicas clássicas, porém muito usadas, conhecidas como métodos de resposta em freqüência. O critério de estabilidade de Nyquist é deduzido e aplicado à análise de estabilidade em sistemas de controle.

O Cap. 10 apresenta técnicas de projeto e compensação. São apresentados exemplos de compensação de sistemas utilizando as técnicas do lugar das raízes e de resposta em freqüência.

O Cap. 11 discute a análise por função-descritiva de sistemas de controle não lineares. O Cap. 12 é relacionado com o bem conhecido método do plano de fase. O Cap. 13 introduz sistemas discretos no tempo e apresenta o método da transformada-z.

Os Caps. 14-16 apresentam uma introdução à teoria de controle moderna. Especificamente, o Cap. 14 introduz o conceito de estado, variáveis de estado e espaço de estado. A representação no espaço de estados de sistemas de controle e a solução de equações de estado são então apresentadas. O Cap. 15 refere-se aos métodos de Liapunov para análise de estabilidade. Finalmente, O Cap. 16 apresenta os conceitos de controlabilidade e observabilidade, sistemas de controle ótimo e uma introdução a sistemas de controle adaptativo.

Os métodos de resposta em freqüência, a parte fundamental da teoria de

controle clássica, são muito usados na indústria. São úteis no tratamento de sistemas de controle invariáveis no tempo com saída-simples-entrada-simples.

A teoria de controle moderna apresenta muitas vantagens em relação à teoria de controle clássica. A primeira é aplicável no projeto de sistemas multivariáveis e variáveis no tempo. Possibilita ao engenheiro de controle levar em conta condições iniciais arbitrárias na síntese de controle ótimo de sistemas. Nesta síntese necessitamos considerar apenas os aspectos analíticos do problema. Um computador digital pode ser programado para operar todos os cálculos numéricos necessários. Esta é uma das vantagens fundamentais da teoria de controle moderna.

É importante salientar que a teoria de controle moderna não substitui completamente a teoria de controle clássica. As duas abordagens complementam-se. A engenharia de controle moderna é baseada nos aspectos úteis das teorias de controle clássica e moderna. A finalidade deste livro é fornecer ao leitor uma boa base e o uso das ferramentas da engenharia de controle moderna.

PROBLEMAS ILUSTRATIVOS E SOLUÇÕES

Problema A.1.1 Liste as principais vantagens e desvantagens de sistemas de controle em malha-aberta.

Solução. As vantagens dos sistemas de controle em malha-aberta são as seguintes:

1. Construção simples e facilidade de manutenção.
2. Menos dispendioso do que um sistema em malha-fechada correspondente.
3. Não há problema de estabilidade.
4. Conveniente quando a saída é difícil de medir ou economicamente não disponível. (Por exemplo, seria bem caro desenvolver um dispositivo para medir a qualidade da saída de uma torradeira.)

As desvantagens de sistemas de controle em malha-aberta são as seguintes:

1. Distúrbios e variações na calibração acarretam erros, e a saída pode ser diferente da desejada.
2. Para manter a qualidade requerida na saída, é necessária a recalibração de tempos em tempos.

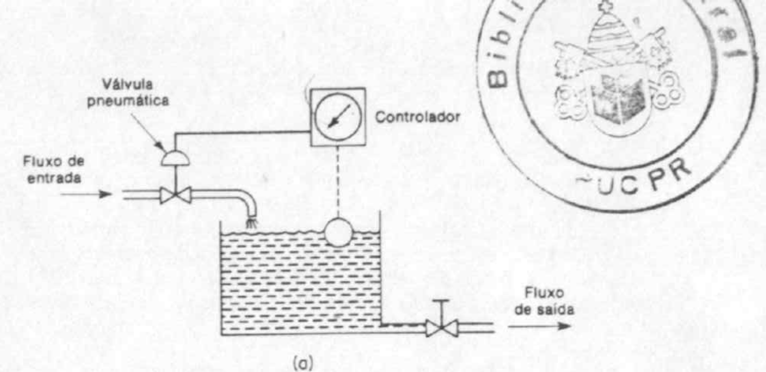
Problema A.1.2 A Fig. 1.10(a) representa um diagrama esquemático de um sistema de controle de nível de líquido. Neste diagrama o controlador automático mantém o nível do líquido comparando o nível real com um nível desejado e corrigindo qualquer erro pelo ajuste da abertura da válvula pneumática. A Fig. 1.10(b) apresenta um diagrama de blocos do sistema de controle. Desenhe o diagrama de blocos correspondente para um sistema de controle de nível de líquido operado por um ser humano.

Solução. No sistema operado por um ser humano, os olhos, o cérebro e os músculos correspondem ao dispositivo de medida, controlador e válvula pneumática, respectivamente. Um diagrama de blocos é indicado na Fig. 1.11.

Problema A.1.3 Um sistema organizacional de engenharia é composto de muitos grupos, tais como administração, pesquisa e desenvolvimento, projeto preliminar, experimentos, projeto do produto e desenhos, fabricação e montagem, e testes. Estes grupos são inter-relacionados a fim de constituírem uma operação global conjunta.

O sistema pode ser analisado pela redução a um conjunto mais elementar de componentes necessários, os quais podem fornecer o detalhamento analítico exigido, bem como por um conjunto de equações simples que representam as características dinâmicas de cada componente. (O desempenho dinâmico de um sistema como este pode ser determinado a partir da relação entre a execução progressiva e o tempo.)

Desenhe um diagrama de blocos funcional mostrando um sistema organizacional de engenharia.



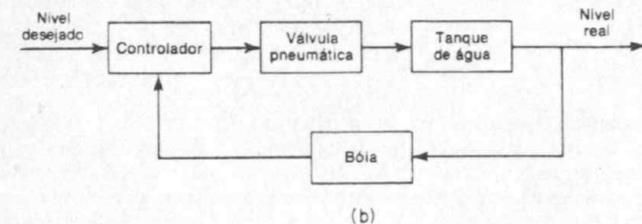


Fig. 1.10 (a) Sistema de controle de nível de líquido; (b) diagrama de blocos.

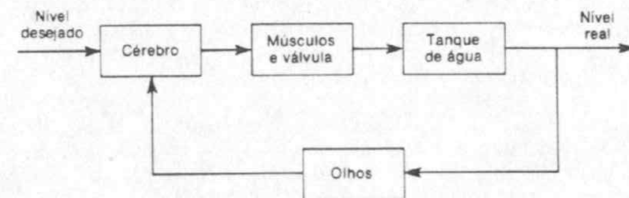


Fig. 1.11 Diagrama de blocos de um sistema de controle de nível de líquido operado por um ser humano.

Solução. Um diagrama de blocos funcional pode ser desenhado utilizando-se blocos para representarem as atividades funcionais e linhas de sinal de interligação para representarem a informação ou saída do produto da operação do sistema. Um diagrama de blocos possível é mostrado na Fig. 1.12.

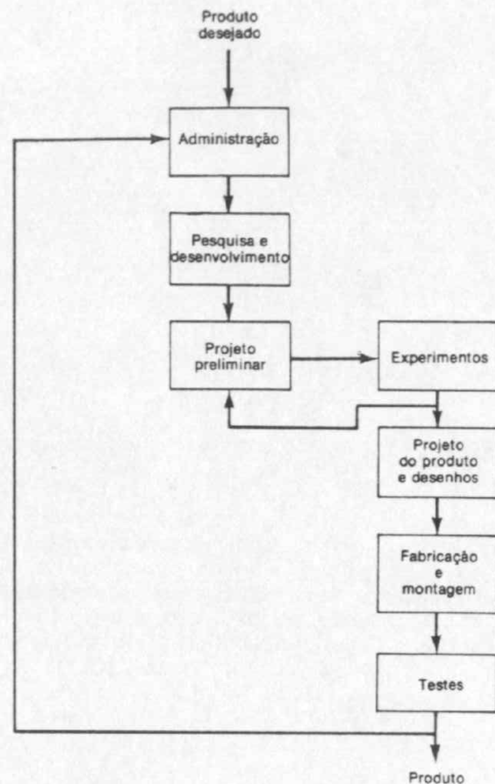


Fig. 1.12 Diagrama de blocos de um sistema organizacional de engenharia.

PROBLEMAS

Problema B.1.1 Muitos sistemas de controle em malha-aberta e em malha-fechada podem ser encontrados em uma residência. Relacione alguns exemplos e descreva-os.

Problema B.1.2 Desenhe um diagrama de blocos de um sistema de aquecimento residencial. Note que um termostato é o controlador do sistema. Estabeleça quais distúrbios poderiam existir em tal sistema.

Problema B.1.3 A Fig. 1.13 indica um sistema de controle de tensão. Explique a seqüência das ações de controle quando a velocidade de alimentação é variada subitamente em um curto intervalo de tempo.

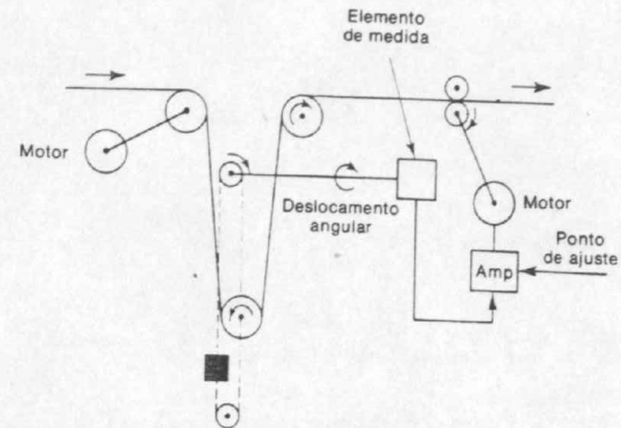


Fig. 1.13 Sistema de controle de tensão.

Problema B.1.4 A Fig. 1.14 mostra um sistema de controle auto-operado. Explique sua operação.

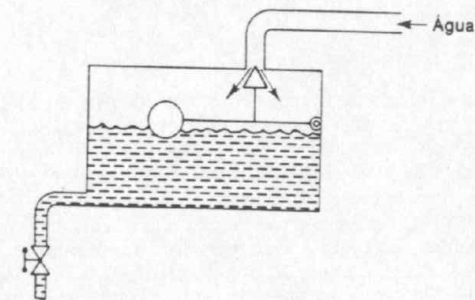


Fig. 1.14 Sistema de controle auto-operado.

2

Base Matemática — Transformada de Laplace

2.1 INTRODUÇÃO

O método da transformada de Laplace é um método operacional que pode ser usado com vantagens para resolver equações diferenciais lineares. Usando transformadas de Laplace, podem-se converter muitas funções comuns, tais como funções senoidais, funções senoidais amortecidas, e funções exponenciais em funções algébricas de uma variável complexa. Operações como diferenciação e integração podem ser substituídas por operações algébricas no plano complexo. Desta forma, uma equação diferencial linear pode ser transformada em uma equação algébrica em uma variável complexa. A solução da equação diferencial pode então ser achada através do uso de uma tabela de transformadas de Laplace ou pelo uso da técnica de expansão em frações parciais, que é apresentada na Seção 2.4.

Uma vantagem do método da transformada de Laplace é que ele permite o uso de técnicas gráficas para prever o desempenho do sistema sem a necessidade de resolver as equações diferenciais deste. Outra vantagem do método da transformada de Laplace é que, quando se resolve a equação diferencial, tanto a componente transitória como a de regime permanente da solução podem ser obtidas simultaneamente.

Revisão de variáveis complexas e funções complexas. Antes de definir a transformada de Laplace da função do tempo $f(t)$, apresentamos uma revisão muito breve de variáveis complexas e funções complexas. Uma variável complexa s tem uma componente real σ e uma componente imaginária $j\omega$, ou $s = \sigma + j\omega$. Uma variável complexa s pode ser representada por um ponto no plano s. A Fig. 2.1 ilustra o plano s e um ponto representativo $s_1 = \sigma_1 + j\omega_1$. Uma função complexa

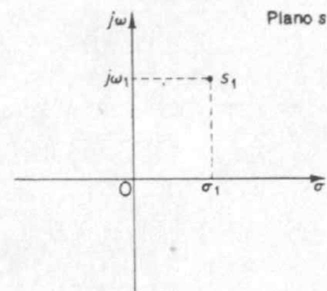


Fig. 2.1 Plano s e um ponto representativo.

$G(s)$, uma função de s, tem uma parte real e uma parte imaginária, ou

$$G(s) = G_x + jG_y$$

onde G_x e G_y são reais. O módulo de uma grandeza complexa $(G_x + jG_y)$ é dado por $\sqrt{G_x^2 + G_y^2}$, e o ângulo θ de $(G_x + jG_y)$ é dado por $\theta = \tan^{-1}(G_y/G_x)$.* A Fig. 2.2 mostra o plano complexo e dois valores complexos representativos. O ângulo θ é medido a partir do eixo real positivo. Uma rotação em sentido anti-horário é definida como a direção positiva para fins de medição de ângulos.

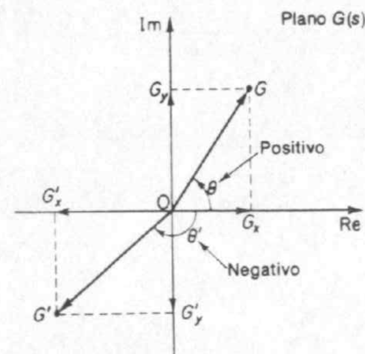


Fig. 2.2 Plano complexo e duas grandezas complexas representativas.

O complexo conjugado de $G(s) = G_x + jG_y$ é definido como $\bar{G}(s) = G_x - jG_y$. Uma quantidade complexa e seu conjugado têm a mesma parte real, mas a parte imaginária de uma é o valor negativo da parte imaginária da outra.

As funções complexas $G(s)$ normalmente encontradas em sistemas de controle linear são funções unívocas de s e são determinadas unicamente para um dado valor

*Usaremos indistintamente a notação arctg ou $\tan^{-1}$ (N. do T.)

de s. Uma função complexa $G(s)$ é dita *analítica* em uma região se $G(s)$ e todas suas derivadas existem naquela região. A derivada de uma função analítica $G(s)$ é dada por

$$\frac{d}{ds}G(s) = \lim_{\Delta s \to 0} \frac{G(s + \Delta s) - G(s)}{\Delta s} = \lim_{\Delta s \to 0} \frac{\Delta G}{\Delta s}$$

O valor da derivada é independente da escolha da trajetória Δs. Uma vez que $\Delta s = \Delta\sigma + j\Delta\omega$, Δs pode se aproximar de zero ao longo de um número infinito de diferentes trajetórias. Pode-se mostrar, embora não seja provado aqui, que se as derivadas tomadas ao longo de duas trajetórias quaisquer, especificamente $\Delta s = \Delta\sigma$ e $\Delta s = j\Delta\omega$, são iguais, então a derivada é unica para qualquer outra trajetória $\Delta s = \Delta\sigma + j\Delta\omega$.

Para uma particular trajetória $\Delta s = \Delta\sigma$ (que significa que a trajetória é paralela ao eixo real),

$$\frac{d}{ds}G(s) = \lim_{\Delta\sigma \to 0}\left(\frac{\Delta G_x}{\Delta\sigma} + j\frac{\Delta G_y}{\Delta\sigma}\right) = \frac{\partial G_x}{\partial\sigma} + j\frac{\partial G_y}{\partial\sigma}$$

Para uma outra trajetória em particular $\Delta s = j\Delta\omega$ (que significa que a trajetória é paralela ao eixo imaginário),

$$\frac{d}{ds}G(s) = \lim_{j\Delta\omega \to 0}\left(\frac{\Delta G_x}{j\Delta\omega} + j\frac{\Delta G_y}{j\Delta\omega}\right) = -j\frac{\partial G_x}{\partial\omega} + \frac{\partial G_y}{\partial\omega}$$

Se estes dois valores da derivada são iguais,

$$\frac{\partial G_x}{\partial\sigma} + j\frac{\partial G_y}{\partial\sigma} = \frac{\partial G_y}{\partial\omega} - j\frac{\partial G_x}{\partial\omega}$$

ou se as duas condições seguintes:

$$\frac{\partial G_x}{\partial\sigma} = \frac{\partial G_y}{\partial\omega} \quad \text{e} \quad \frac{\partial G_y}{\partial\sigma} = -\frac{\partial G_x}{\partial\omega}$$

são satisfeitas, então a derivada $dG(s)/ds$ é determinada unicamente. Estas duas condições são conhecidas como condições de Cauchy-Riemann.

Como um exemplo, considere o seguinte $G(s)$:

$$G(s) = \frac{1}{s+1}$$

Então,

$$G(\sigma + j\omega) = \frac{1}{\sigma + j\omega + 1} = G_x + jG_y$$

onde

$$G_x = \frac{\sigma + 1}{(\sigma + 1)^2 + \omega^2} \quad \text{e} \quad G_y = \frac{-\omega}{(\sigma + 1)^2 + \omega^2}$$

Percebe-se que, exceto no ponto $s = -1$ (explicitamente, $\sigma = -1$, $\omega = 0$), $G(s)$ satisfaz as condições de Cauchy-Riemann:

$$\frac{\partial G_x}{\partial\sigma} = \frac{\partial G_y}{\partial\omega} = \frac{\omega^2 - (\sigma + 1)^2}{[(\sigma + 1)^2 + \omega^2]^2}$$

$$\frac{\partial G_y}{\partial\sigma} = -\frac{\partial G_x}{\partial\omega} = \frac{2\omega(\sigma + 1)}{[(\sigma + 1)^2 + \omega^2]^2}$$

Portanto $G(s) = 1/(s + 1)$ é analítica em todo plano s exceto em $s = -1$. A derivada $dG(s)/ds$, exceto em $s = -1$, é

$$\frac{d}{ds}G(s) = \frac{\partial G_x}{\partial\sigma} + j\frac{\partial G_y}{\partial\sigma} = \frac{\partial G_y}{\partial\omega} - j\frac{\partial G_x}{\partial\omega} = -\frac{1}{(\sigma + j\omega + 1)^2} = -\frac{1}{(s + 1)^2}$$

Note que a derivada de uma função analítica pode ser obtida simplesmente pela diferenciação de $G(s)$ com relação a s. Neste exemplo,

$$\frac{d}{ds}\left(\frac{1}{s+1}\right) = -\frac{1}{(s+1)^2}$$

Pontos do plano s em que a função $G(s)$ é analítica são chamados de pontos *ordinários*, enquanto que pontos do plano s em que a função $G(s)$ não é analítica são chamados de pontos *singulares*. Pontos singulares em que a função $G(s)$ ou suas derivadas se aproximam de infinito são chamados *pólos*. Por exemplo,

$$G(s) = \frac{K(s + z)}{(s + p_1)(s + p_2)^2} \tag{2-1}$$

tem pólos em $s = -p_1$ e $s = -p_2$.

Se $G(s)$ tende a infinito quando s tende a $-p$ e se a função

$$G(s)(s + p)^n \qquad (n = 1, 2, 3, \ldots)$$

tem um valor finito não nulo em $s = -p$, então $s = -p$ é chamado um pólo de ordem n. Se $n = 1$, o pólo é chamado um pólo simples. Se $n = 2, 3, \ldots$, o pólo é chamado de pólo de segunda ordem, pólo de terceira ordem etc.

Pontos em que a função $G(s)$ é igual a zero são chamados de zeros. A função dada pela Eq. (2.1) tem um zero em $s = -z$. Se pontos no infinito são também contados, $G(s)$ tem o mesmo número de pólos que zeros. A função $G(s)$ dada pela Eq. (2.1) tem dois zeros no infinito em adição ao zero finito em $s = -z$, pois

$$\lim_{s\to\infty} G(s) \doteq \lim_{s\to\infty} \frac{K}{s^2} \longrightarrow 0$$

Portanto, $G(s)$ tem três pólos e três zeros (um zero finito e dois infinitos).

Mapeamento conforme. Um mapeamento que preserva tanto o tamanho como o sentido dos ângulos é chamado *conforme*. Considere uma função analítica $z = F(s)$. A relação funcional $z = F(s)$ pode ser interpretada como um mapeamento de pontos do plano s sobre pontos do plano z ou o plano $F(s)$. Para qualquer ponto P do plano s em que $F(s)$ é regular (significando não-singular), corresponde um ponto P' no plano $F(s)$. P' é chamado de imagem de P, sendo dado pela função $z = F(s)$.

Um mapeamento dado por uma função analítica é conforme. Isto é, duas curvas suaves no plano s que se interceptam e formam um ângulo θ são mapeadas sobre duas curvas suaves no plano $F(s)$ que se interceptam e formam o mesmo ângulo θ. (Note que pode ocorrer distorção da forma das duas curvas no mapeamento, embora os ângulos e o sentido dos mesmos sejam preservados.)

Se a função $G(s)$ é regular e unívoca em um domínio Γ, então a imagem de uma curva contínua em Γ mapeada por $z = F(s)$ também é uma curva contínua.

Para provar que o mapeamento associado com uma função analítica $z = F(s)$ é conforme, considere uma curva suave $s = s(\xi)$, que passa por um ponto ordinário s_0. Se escrevermos $z_0 = F(s_0)$, então

$$z - z_0 = \frac{F(s) - F(s_0)}{s - s_0}(s - s_0)$$

Portanto,

$$\angle z - z_0 = \angle \frac{F(s) - F(s_0)}{s - s_0} + \angle s - s_0$$

onde $\angle s - s_0$ é o ângulo entre o eixo real positivo e o vetor apontando de s_0 a s. Se s tende a s_0 ao longo da curva suave $s(\xi)$, então $\angle s - s_0$ é o ângulo θ_1 entre o eixo real positivo e a tangente à curva em s_0. Da mesma forma, à medida que z tende a z_0, $\angle z - z_0$ tende ao ângulo ϕ_1, que é o ângulo entre o eixo real positivo e a tangente a $F(s)$ em z_0. Portanto, obtemos

$$\phi_1 = \angle F'(s_0) + \theta_1$$

contanto que $F'(s_0)$, a derivada de $F(s)$ calculada em $s = s_0$ não é zero. (Note que a menos que $F'(s_0) \neq 0$, $\angle F'(s_0)$ não pode ser definida.) Portanto

$$\phi_1 - \theta_1 = \angle F'(s_0)$$

Observe também que $\phi_1 - \theta_1$ depende não da curva suave $s = s(\xi)$ escolhida em particular, mas sim do ponto s_0.

Usando uma outra curva suave $s = s_2(\xi)$ passando pelo ponto s_0, podemos fazer uma análise similar, obtendo

$$\phi_2 - \theta_2 = \angle F'(s_0)$$

Portanto

$$\phi_1 - \theta_1 = \phi_2 - \theta_2$$

ou

$$\phi_2 - \phi_1 = \theta_2 - \theta_1$$

Portanto, o tamanho e sentido dos ângulos são preservados neste mapeamento. Acabamos então de ver que o mapeamento dado pela função analítica $z = F(s)$ é conforme em todos pontos em que $F(s)$ é regular e $F'(s) \neq 0$.

Utilizaremos mapeamento conforme ao discutir os gráficos do método do lugar das raízes (Cap. 8) e o critério de estabilidade de Nyquist (Cap. 9).

Finalmente, observe que pode ser provado que um mapeamento não pode ser conforme em todos pontos de um domínio Γ a não ser que o mapeamento seja dado por uma função analítica $z = F(s)$ que é regular no domínio. A conformidade de mapeamento é uma característica de funções analíticas.

2.2 TRANSFORMADA DE LAPLACE

Esta seção apresenta uma definição da transformada de Laplace, uma breve discussão da condição para a existência da transformada de Laplace, e exemplos para ilustrar a obtenção das transformadas de Laplace de várias funções comumente utilizadas.

Vamos definir

$f(t)$ = uma função do tempo t tal que $f(t) = 0$ para $t < 0$
s = uma variável complexa
$\mathscr{L}$ = um símbolo operacional indicando que a quantidade que ele prefixa é para ser transformada pela integral de Laplace $\int_0^\infty e^{-st} dt$
$F(s)$ = transformada de Laplace de $f(t)$

Então a transformada de Laplace de $f(t)$ é definida por

$$\mathscr{L}[f(t)] = F(s) = \int_0^\infty e^{-st}\, dt[f(t)] = \int_0^\infty f(t)\, e^{-st}\, dt$$

Exemplo 2.1 Função exponencial. Considere a seguinte função:

$$f(t) = 0 \quad \text{para } t < 0$$
$$= Ae^{-\alpha t} \quad \text{para } t \geq 0$$

onde A e α são constantes. A transformada de Laplace de $f(t)$ é obtida da seguinte forma:

$$\begin{aligned}\mathcal{L}[f(t)] &= \int_0^\infty Ae^{-\alpha t}e^{-st}\,dt \\ &= A\int_0^\infty e^{-(\alpha+s)t}\,dt \\ &= \frac{A}{s+\alpha}\end{aligned}$$

Vemos então que a função exponencial produz um pólo no plano complexo. Ao realizar esta integração, supusemos que a parte real de s era maior que $-\alpha$. (Observe que tal suposição é necessária para tornar a integral absolutamente convergente.)

A transformada de Laplace de uma função $f(t)$ existe se $f(t)$ é seccionalmente contínua em todo intervalo finito na região $t > 0$ e se a função é de ordem exponencial quando t tende a infinito. Em outras palavras, a integral de Laplace deve convergir. Uma função $f(t)$ é de ordem exponencial se existe uma constante σ real e positiva tal que a função

$$e^{-\sigma t}\,|f(t)|$$

tende a zero quando t tende a infinito. Se o limite da função $e^{-\sigma t}|f(t)|$ tende a zero para σ maior que σ_c e o limite tende a infinito para σ menor que σ_c, o valor σ_c é chamado de *abscissa de convergência*.

Para a função $f(t) = Ae^{-\alpha t}$

$$\lim_{t\to\infty} e^{-\sigma t}\,|Ae^{-\alpha t}|$$

tende a zero se $\sigma > -\alpha$. A abscissa de convergência neste caso é $\sigma_c = -\alpha$. A integral $\int_0^\infty f(t)e^{-st}dt$ converge apenas se σ, a parte real de s, é maior que a abscissa de convergência σ_c. Portanto o operador s deve ser escolhido como uma constante tal que esta integral convirja.

Em termos dos pólos da função $F(s)$, a abscissa de convergência σ_c corresponde à parte real do pólo localizado mais distante à direita no plano s. Por exemplo, para a seguinte função $F(s)$,

$$F(s) = \frac{K(s+3)}{(s+1)(s+2)}$$

a abscissa de convergência σ_c é igual a -1. Pode-se ver que para funções como t, sen ωt, e t sen ωt a abscissa de convergência é igual a zero. Para funções como e^{-ct}, te^{-ct}, e^{-ct} sen ωt etc., a abscissa de convergência é igual a $-c$. Para funções que crescem mais depressa que a função exponencial é impossível se achar valores adequados para a abscissa de convergência. Portanto, funções como e^{t^2} e te^{t^2} não possuem transformadas de Laplace.

O leitor deve ser alertado para o fato de que, embora e^{t^2} (para $0 \le t \le \infty$) não possua uma transformada de Laplace, a função do tempo definida por

$$\begin{aligned}f(t) &= e^{t^2} \quad \text{para } 0 \le t \le T < \infty \\ &= 0 \quad \text{para } t < 0,\ T < t\end{aligned}$$

possui uma transformada de Laplace pois $f(t) = e^{t^2}$ em apenas um intervalo limitado $0 \le t \le T$ e não em $0 \le t \le \infty$. Tal tipo de sinal pode ser gerado fisicamente. Observe que os sinais que podemos gerar fisicamente sempre têm uma correspondente transformada de Laplace.

Ao derivar a transformada de Laplace de $f(t) = Ae^{-\alpha t}$ no Exemplo 2.1, impusemos que a parte real de s fosse maior que $-\alpha$ (a abscissa de convergência). Uma pergunta que pode surgir de imediato é se a transformada de Laplace obtida é válida ou não na região $\sigma < -\alpha$ do plano s. Para responder a esta pergunta, devemos utilizar a teoria de variáveis complexas. Na teoria de variáveis complexas, há um teorema conhecido como o teorema de extensão analítica. Este diz que se duas funções analíticas são iguais em um comprimento finito ao longo de qualquer arco em uma região em que ambas são analíticas, então elas são iguais em todo lugar desta região. O arco de igualdade é normalmente o eixo real, ou uma porção deste. Usando este teorema, a forma de $F(s)$, determinada por uma integração em que s pode tomar qualquer valor real positivo maior que a abscissa de convergência, vale para quaisquer valores complexos de s em que $F(s)$ é analítica. Portanto, embora façamos a exigência da parte real de s ser maior que a abscissa de convergência para fazer a integral $\int_0^\infty f(t)e^{-st}dt$ ser absolutamente convergente, uma vez que a transmada de Laplace é obtida, $F(s)$ pode ser considerada válida ao longo de todo o plano s exceto nos pólos de $F(s)$.

Se uma função $f(t)$ tem uma transformada de Laplace, então a transformada de Laplace de $Af(t)$, onde A é uma constante, é dada por

$$\mathcal{L}[Af(t)] = A\mathcal{L}[f(t)]$$

Isto é óbvio pela definição da transformada de Laplace. Da mesma forma, se as funções $f_1(t)$ e $f_2(t)$ têm transformadas de Laplace, então a transformada de Laplace da função $f_1(t) + f_2(t)$ é dada por

$$\mathcal{L}[f_1(t) + f_2(t)] = \mathcal{L}[f_1(t)] + \mathcal{L}[f_2(t)]$$

Novamente, a demonstração desta igualdade é evidente da definição da transformada de Laplace.

A seguir, derivamos transformadas de Laplace de algumas funções encontradas com freqüência. Note que a transformada de Laplace de qualquer função $f(t)$ transformável segundo Laplace é obtida multiplicando-se $f(t)$ por e^{-st} e então integrando o produto de $t = 0$ a $t = \infty$. (Note também que uma vez que sabemos o método de obtenção da transformada de Laplace, não é necessário derivar a transformada de Laplace de $f(t)$ a cada vez. Podem-se utilizar tabelas de transformada de Laplace para achar a transformada de uma dada função $f(t)$.)

Exemplo 2.2 Função degrau. Considere a seguinte função degrau:

$$\begin{aligned}f(t) &= 0 \quad &\text{para } t < 0 \\ &= A = \text{constante} \quad &\text{para } t > 0\end{aligned}$$

A função degrau aqui não é definida em $t = 0$. Mas isto é irrelevante pois

$$\int_{0-}^{0+} Ae^{-st}\,dt = 0$$

A transformada de Laplace de $f(t)$ é dada por

$$\mathscr{L}[f(t)] = \int_0^\infty Ae^{-st}\,dt = \frac{A}{s}$$

Ao fazer esta integração, supomos que a parte real de s é maior que zero (a abscissa de convergência) e portanto que $\lim_{t\to\infty} e^{-st}$ é zero. Como explicado anteriormente, a transformada de Laplace é válida em todo plano s exceto no pólo $s = 0$.

A função degrau cuja amplitude é unitária é chamada função *degrau unitário*. A função degrau unitário que ocorre em $t = t_0$ é freqüentemente escrita como $u(t - t_0)$ ou $1(t - t_0)$. Neste livro usaremos a notação $1(t - t_0)$ a não ser que haja menção em contrário. A função degrau de amplitude A pode ser escrita como $f(t) = A\,1(t)$. A transformada de Laplace da função degrau unitário que é definida por

$$\begin{aligned} 1(t) &= 0 \quad \text{para } t < 0 \\ &= 1 \quad \text{para } t > 0 \end{aligned}$$

é $1/s$, ou

$$\mathscr{L}[1(t)] = \frac{1}{s}$$

Fisicamente, uma função degrau ocorrendo em $t = 0$ corresponde a um sinal constante aplicado subitamente ao sistema no instante t igual a zero.

Exemplo 2.3 Função rampa. Considere a seguinte função rampa:

$$\begin{aligned} f(t) &= 0 \quad \text{para } t < 0 \\ &= At \quad \text{para } t \geq 0 \end{aligned}$$

onde A é uma constante. A transformada de Laplace da função rampa é dada por

$$\begin{aligned} \mathscr{L}[f(t)] &= A\int_0^\infty te^{-st}\,dt \\ &= At\,\frac{e^{-st}}{-s}\Big|_0^\infty - \int_0^\infty \frac{Ae^{-st}}{-s}\,dt \\ &= \frac{A}{s}\int_0^\infty e^{-st}\,dt \\ &= \frac{A}{s^2} \end{aligned}$$

Exemplo 2.4 Função senoidal. A transformada de Laplace da seguinte função senoidal:

$$\begin{aligned} f(t) &= 0 \quad \text{para } t < 0 \\ &= A\operatorname{sen}\omega t \quad \text{para } t \geq 0 \end{aligned}$$

onde A e ω são constantes, é obtida da seguinte forma:

$$\mathscr{L}[f(t)] = A\int_0^\infty (\operatorname{sen}\omega t)e^{-st}\,dt$$

Uma vez que

$$\begin{aligned} e^{j\omega t} &= \cos\omega t + j\operatorname{sen}\omega t \\ e^{-j\omega t} &= \cos\omega t - j\operatorname{sen}\omega t \end{aligned}$$

obtemos

$$\operatorname{sen}\omega t = \frac{1}{2j}(e^{j\omega t} - e^{-j\omega t})$$

Portanto,

$$\begin{aligned} \mathscr{L}[f(t)] &= \frac{A}{2j}\int_0^\infty (e^{j\omega t} - e^{-j\omega t})e^{-st}\,dt \\ &= \frac{A}{2j}\frac{1}{s - j\omega} - \frac{A}{2j}\frac{1}{s + j\omega} \\ &= \frac{A\omega}{s^2 + \omega^2} \end{aligned}$$

A transformada de Laplace de $\cos\omega t$ pode ser obtida de forma semelhante.

2.3 TEOREMAS DA TRANSFORMADA DE LAPLACE

Nesta seção apresentamos vários teoremas da transformada de Laplace que são úteis no estudo de sistemas de controle linear.

Função transladada. Obteremos a transformada de Laplace da função transladada $f(t - \alpha)$. Aqui supomos que $f(t)$ é zero para $t < 0$ ou $f(t - \alpha) = 0$ para $t < \alpha$. As funções $f(t)$ e $f(t - \alpha)$ são mostradas graficamente na Fig. 2.3.

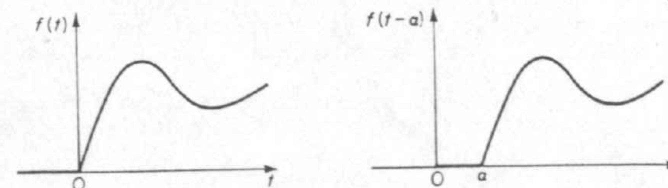


Fig. 2.3 Gráficos de $f(t)$ e $f(t - \alpha)$.

Como $f(t-\alpha) = 0$ para $0 < t < \alpha$,

$$F(s) = \int_0^\infty f(\tau)e^{-s\tau}\,d\tau$$
$$= \int_\alpha^\infty f(t-\alpha)e^{-s(t-\alpha)}\,dt$$
$$= e^{\alpha s}\int_0^\infty f(t-\alpha)e^{-st}\,dt$$

onde $t - \alpha = \tau$. Portanto,

$$\mathscr{L}[f(t-\alpha)] = \int_0^\infty f(t-\alpha)e^{-st}\,dt = e^{-\alpha s}F(s)$$

Esta última equação diz que a translação da função do tempo $f(t)$ de α unidades de tempo corresponde à multiplicação da transformada $F(s)$ por $e^{-\alpha s}$.

Para mostrar claramente a implicação de que $f(t-\alpha)$ é zero para $t < \alpha$, é necessário escrever a função transladada como $f(t-\alpha)\,1(t-\alpha)$. Com esta notação,

$$\mathscr{L}[f(t-\alpha)\,1(t-\alpha)] = e^{-\alpha s}F(s)$$

Exemplo 2.5 Função pulso. Considere a seguinte função pulso:

$$f(t) = A = \text{constante} \quad \text{para } 0 < t < t_0$$
$$= 0 \quad \text{para } t < 0,\ t_0 < t$$

A transformada de Laplace de $f(t)$ é obtida da seguinte forma: A função pulso $f(t)$ pode ser considerada uma função degrau de amplitude A que começa em $t = 0$ e que é superposta a um degrau negativo de altura A começando em $t = t_0$; explicitamente,

$$f(t) = A1(t) - A1(t - t_0)$$

A transformada de Laplace de $f(t)$ é então obtida como

$$\mathscr{L}[f(t)] = \mathscr{L}[A1(t)] - \mathscr{L}[A1(t-t_0)]$$
$$= \frac{A}{s} - \frac{A}{s}e^{-st_0}$$
$$= \frac{A}{s}(1 - e^{-st_0})$$

Exemplo 2.6 Função impulso. A função impulso é um caso limite especial de uma função pulso. Considere a seguinte função impulso:

$$f(t) = \lim_{t_0\to 0}\frac{A}{t_0} \quad \text{para } 0 < t < t_0$$
$$= 0 \quad \text{para } t < 0,\ t_0 < t$$

Como a amplitude da função impulso é A/t_0 e a duração é t_0, a área sob o impulso é igual a A. Como a duração t_0 tende a zero, a altura A/t_0 tende a infinito, mas a área sob o impulso permanece igual a A. Note que o tamanho de um impulso é medido pela sua área.

A transformada de Laplace da função impulso $f(t)$ pode ser obtida como segue:

$$\mathscr{L}[f(t)] = \lim_{t_0\to 0}\frac{A}{t_0 s}(1 - e^{-st_0})$$
$$= \lim_{t_0\to 0}\frac{\frac{d}{dt_0}[A(1-e^{-st_0})]}{\frac{d}{dt_0}(t_0 s)}$$
$$= \frac{As}{s}$$
$$= A$$

Portanto, a transformada de Laplace da função impulso é igual à área sob o impulso.

A função impulso cuja área é igual à unidade é chamada função *impulso unitário* ou função delta de Dirac. A função impulso unitário, ocorrendo em $t = t_0$, é normalmente indicada por $\delta(t - t_0)$. $\delta(t - t_0)$ satisfaz as seguintes condições:

$$\delta(t - t_0) = 0 \quad \text{para } t \neq t_0$$
$$\delta(t - t_0) = \infty \quad \text{para } t = t_0$$
$$\int_{-\infty}^{\infty}\delta(t - t_0)\,dt = 1$$

Um impulso que tem amplitude infinita e duração nula é uma ficção matemática e não ocorre em sistemas físicos. Se, entretanto, a amplitude de um pulso de entrada a um sistema é muito grande e sua duração muito pequena comparada com as constantes de tempo do sistema, então podemos aproximar a entrada em pulso por uma entrada em função impulso. Por exemplo, se uma força ou torque de entrada $f(t)$ é aplicada a um sistema por uma duração muito curta $0 < t < t_0$, com a magnitude de $f(t)$ sendo suficientemente grande de tal forma a tornar a integral $\int_0^{t_0} f(t)\,dt$ não desprezível, então esta entrada pode ser considerada uma entrada impulsiva. A entrada impulsiva fornece energia ao sistema em um tempo infinitesimal.

Deve-se observar que quando descrevemos a entrada impulsiva, a magnitude ou tamanho do impulso é de máxima importância, mas a forma do impulso normalmente é irrelevante. De fato, qualquer pulso de duração muito curta pode ser considerado um impulso unitário se ele satisfaz a condição de que a área embaixo da curva permanece unitária quando a largura do pulso tende a zero.

O conceito da função impulsiva é bastante útil na diferenciação de funções descontínuas. A função impulso unitário $\delta(t)$ pode ser considerada como a derivada em função do tempo da função degrau unitário $1(t)$ no ponto de descontinuidade, ou

$$\delta(t) = \frac{d}{dt}1(t)$$

Ou inversamente, se a função impulso unitário $\delta(t)$ é integrada, o resultado é a função degrau unitário $1(t)$. Com o conceito da função impulso, podemos diferenciar uma função contendo descontinuidades, obtendo impulsos, cujas magnitudes são iguais à magnitude de cada descontinuidade correspondente.

Multiplicação de $f(t)$ por $e^{-\alpha t}$. Se $f(t)$ é transformável por Laplace, com sua transformada de Laplace sendo $F(s)$, então a transformada de Laplace de $e^{-\alpha t}f(t)$ é obtida como segue:

$$\begin{aligned}\mathscr{L}[e^{-\alpha t}f(t)] &= \int_0^\infty e^{-\alpha t}f(t)e^{-st}\,dt \\ &= F(s+\alpha) \qquad (2\text{-}2)\end{aligned}$$

Vemos que a multiplicação de $f(t)$ por $e^{-\alpha t}$ tem o efeito de substituir s por $(s+\alpha)$ na transformada de Laplace. Ou, inversamente, mudando s para $(s+\alpha)$ é equivalente a multiplicar $f(t)$ por $e^{-\alpha t}$. (Note que α pode ser real ou complexo.)

A relação dada pela Eq. (2.2) é muito útil para se determinar transformadas de Laplace de funções como $e^{-\alpha t}$ sen ωt e $e^{-\alpha t}\cos\omega t$. Por exemplo, como

$$\mathscr{L}[\text{sen}\,\omega t] = \frac{\omega}{s^2+\omega^2} = F(s)$$

então, da Eq. (2.2), segue-se que a transformada de Laplace de $e^{-\alpha t}$ sen ωt é dada por

$$\begin{aligned}\mathscr{L}[e^{-\alpha t}\,\text{sen}\,\omega t] &= F(s+\alpha) \\ &= \frac{\omega}{(s+\alpha)^2+\omega^2}\end{aligned}$$

Mudança de escala de tempo. Ao analisar sistemas físicos, às vezes se deseja mudar a escala de tempo ou normalizar uma dada função do tempo. O resultado obtido em termos de tempo normalizado é útil porque pode ser aplicado diretamente a diferentes sistemas tendo equações matemáticas similares.

Se t é modificado para t/α, onde α é uma constante positiva, então a função $f(t)$ é modificada para $f(t/\alpha)$. Se denotarmos a transformada de Laplace de $f(t)$ por $F(s)$, então a transformada de Laplace de $f(t/\alpha)$ pode ser obtida como visto abaixo:

$$\mathscr{L}\left[f\left(\frac{t}{\alpha}\right)\right] = \int_0^\infty f\left(\frac{t}{\alpha}\right)e^{-st}\,dt$$

Fazendo $t/\alpha = t_1$ e $\alpha s = s_1$, obtemos

$$\begin{aligned}\mathscr{L}\left[f\left(\frac{t}{\alpha}\right)\right] &= \int_0^\infty f(t_1)e^{-s_1t_1}\,d(\alpha t_1) \\ &= \alpha\int_0^\infty f(t_1)e^{-s_1t_1}\,dt_1 \\ &= \alpha F(s_1)\end{aligned}$$

ou

$$\mathscr{L}\left[f\left(\frac{t}{\alpha}\right)\right] = \alpha F(\alpha s)$$

Como um exemplo, considere $f(t) = e^{-t}$ e $f(t/5) = e^{-0,2t}$. Obtemos

$$\mathscr{L}[f(t)] = \mathscr{L}[e^{-t}] = F(s) = \frac{1}{s+1}$$

Portanto

$$\mathscr{L}\left[f\left(\frac{t}{5}\right)\right] = \mathscr{L}[e^{-0,2t}] = 5F(5s) = \frac{5}{5s+1}$$

Este resultado pode ser verificado facilmente tomando-se a transformada de Laplace de $e^{-0,2t}$ diretamente como segue:

$$\mathscr{L}[e^{-0,2t}] = \frac{1}{s+0,2} = \frac{5}{5s+1}$$

Observação sobre o limite inferior da integral de Laplace. Em alguns casos, $f(t)$ possui uma função impulso em $t = 0$. Então o limite inferior da integral de Laplace deve ser claramente especificado, se é 0_- ou 0_+, pois as transformadas de Laplace de $f(t)$ diferem para estes dois limites inferiores. Se tal distinção do limite inferior da integral de Laplace for necessária, então usaremos as seguintes notações:

$$\begin{aligned}\mathscr{L}_+[f(t)] &= \int_{0+}^\infty f(t)e^{-st}\,dt \\ \mathscr{L}_-[f(t)] &= \int_{0-}^\infty f(t)e^{-st}\,dt = \mathscr{L}_+[f(t)] + \int_{0-}^{0+} f(t)e^{-st}\,dt\end{aligned}$$

Se $f(t)$ envolve uma função impulsiva em $t = 0$, então

$$\mathscr{L}_+[f(t)] \neq \mathscr{L}_-[f(t)]$$

pois

$$\int_{0-}^{0+} f(t)e^{-st}\,dt \neq 0$$

para tal caso. Obviamente, se $f(t)$ não possui uma função impulso em $t = 0$,

$$\mathscr{L}_+[f(t)] = \mathscr{L}_-[f(t)]$$

Teorema da diferenciação. A transformada de Laplace de uma função $f(t)$ é dada por

$$\mathscr{L}\left[\frac{d}{dt}f(t)\right] = sF(s) - f(0) \qquad (2\text{-}3)$$

onde $f(0)$ é o valor inicial de $f(t)$, calculado em $t = 0$.

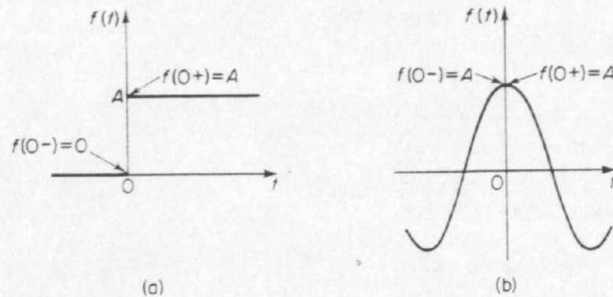


Fig. 2.4 (a) Função degrau; (b) função co-seno.

Para uma dada função $f(t)$, os valores de $f(0_+)$ e $f(0_-)$ podem ser iguais ou diferentes, como ilustrado na Fig. 2.4. A distinção entre $f(0_+)$ e $f(0_-)$ é importante quando $f(t)$ tem uma descontinuidade em $t = 0$, porque em tal caso $df(t)/dt$ vai ter uma função impulso em $t = 0$. Se $f(0_-) = f(0_-)$, a Eq. (2.3) deve ser modificada para

$$\mathscr{L}_+\left[\frac{d}{dt}f(t)\right] = sF(s) - f(0+)$$

$$\mathscr{L}_-\left[\frac{d}{dt}f(t)\right] = sF(s) - f(0-)$$

Para demonstrar o teorema da diferenciação, procedemos como segue: Integrando a integral de Laplace por partes, obtemos

$$\int_0^\infty f(t)e^{-st}\,dt = f(t)\frac{e^{-st}}{-s}\Big|_0^\infty - \int_0^\infty \left[\frac{d}{dt}f(t)\right]\frac{e^{-st}}{-s}\,dt$$

Portanto,

$$F(s) = \frac{f(0)}{s} + \frac{1}{s}\mathscr{L}\left[\frac{d}{dt}f(t)\right]$$

Resultando

$$\mathscr{L}\left[\frac{d}{dt}f(t)\right] = sF(s) - f(0)$$

Da mesma forma obtemos as seguintes relações para a segunda derivada de $f(t)$:

$$\mathscr{L}\left[\frac{d^2}{dt^2}f(t)\right] = s^2F(s) - sf(0) - \dot{f}(0)$$

onde $\dot{f}(0)$ é o valor de $df(t)/dt$ calculado em $t = 0$. Para deduzir esta equação vamos definir

$$\frac{d}{dt}f(t) = g(t)$$

Então

$$\begin{aligned}\mathscr{L}\left[\frac{d^2}{dt^2}f(t)\right] &= \mathscr{L}\left[\frac{d}{dt}g(t)\right]\\ &= s\mathscr{L}[g(t)] - g(0)\\ &= s\mathscr{L}\left[\frac{d}{dt}f(t)\right] - \dot{f}(0)\\ &= s^2F(s) - sf(0) - \dot{f}(0)\end{aligned}$$

Similarmente, para a derivada n-ésima de $f(t)$, obtemos

$$\mathscr{L}\left[\frac{d^n}{dt^n}f(t)\right] = s^nF(s) - s^{n-1}f(0) - s^{n-2}\dot{f}(0) - \cdots - s\overset{(n-2)}{f}(0) - \overset{(n-1)}{f}(0)$$

onde $f(0), \dot{f}(0), \ldots, \overset{(n-1)}{f}(0)$ representam os valores de $f(t)$, $df(t)/dt$, $\ldots d^{n-1}f(t)/dt^{n-1}$, respectivamente, calculados em $t = 0$. Se a distinção entre $\mathscr{L}_+$ e $\mathscr{L}_-$ é necessária, substituímos $t = 0_+$ ou $t = 0_-$ em $f(t), \dot{f}(t), \ldots, \overset{(n-1)}{f}(t)$, dependendo se tomamos $\mathscr{L}_+$ ou $\mathscr{L}_-$.

Note que, para a transformada de Laplace da derivada n-ésima de $f(t)$ existir, $d^nf(t)/dt^n$ deve ser transformável segundo Laplace.

Note também que se todos valores iniciais de $f(t)$ e suas derivadas são iguais a zero, então a transformada de Laplace da n-ésima derivada de $f(t)$ é dada por $s^nF(s)$.

Exemplo 2.7 Considere a seguinte função co-seno:

$$\begin{aligned} g(t) &= 0 && \text{para } t < 0\\ &= A\cos\omega t && \text{para } t \geq 0\end{aligned}$$

A transformada de Laplace da função co-seno pode ser obtida diretamente como no caso da função senoidal. O uso do teorema da diferenciação será mostrado aqui, entretanto, derivando a transformada de Laplace da função co-seno a partir da transformada de Laplace da função seno. Se definirmos

$$\begin{aligned} f(t) &= 0 && \text{para } t < 0\\ &= \text{sen}\,\omega t && \text{para } t \geq 0\end{aligned}$$

Então

$$F(s) = \mathscr{L}[\text{sen}\,\omega t] = \frac{\omega}{s^2 + \omega^2}$$

A transformada de Laplace da função co-seno é obtida como

$$\begin{aligned}\mathscr{L}[A\cos\omega t] &= \mathscr{L}\left[\frac{d}{dt}\frac{A}{\omega}\operatorname{sen}\omega t\right] \\ &= \frac{A}{\omega}[sF(s) - f(0)] \\ &= \frac{A}{\omega}\left(\frac{s\omega}{s^2+\omega^2} - 0\right) \\ &= \frac{As}{s^2+\omega^2}\end{aligned}$$

Teorema do valor final. Se $f(t)$ e $df(t)/dt$ são transformáveis segundo Laplace, se $\lim_{t\to\infty} f(t)$ existe, e se $sF(s)$ é analítica no semiplano direito do plano s incluindo o eixo $j\omega$, exceto por um pólo simples na origem (que significa que $f(t)$ tende para um valor bem definido quando $t \to \infty$), então

$$\lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s)$$

Para provar o teorema, fazemos s tender a zero na equação para a transformada de Laplace da derivada de $f(t)$, ou

$$\lim_{s\to 0}\int_0^\infty \left[\frac{d}{dt}f(t)\right]e^{-st}\,dt = \lim_{s\to 0}[sF(s) - f(0)]$$

Como $\lim_{s\to 0} e^{-st} = 1$, obtemos

$$\begin{aligned}\int_0^\infty \left[\frac{d}{dt}f(t)\right]dt = f(t)\Big|_0^\infty &= f(\infty) - f(0) \\ &= \lim_{s\to 0} sF(s) - f(0)\end{aligned}$$

Portanto,

$$f(\infty) = \lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s)$$

O teorema do valor final diz que o comportamento em regime estacionário de $f(t)$ é o mesmo que o comportamento de $sF(s)$ na vizinhança de $s = 0$. Portanto, é possível obter o valor de $f(t)$ em $t = \infty$ diretamente de $F(s)$.

Percebe-se que quando $f(t)$ é a função senoidal sen ωt, $sF(s)$ tem pólos em $s = \pm j\omega$ e $\lim_{t\to\infty} f(t)$ não existe. Portanto, este teorema não é válido para tal função. Se $f(t)$ tende a infinito quando t tende a infinito, então $\lim_{t\to\infty} f(t)$ não existe, e o teorema do valor final não se aplica a este caso. Devemos nos assegurar que todas as condições do teorema do valor final são satisfeitas antes de aplicá-lo para um dado problema.

O teorema do valor inicial, que segue adiante, é o oposto do teorema do valor final. O teorema do valor inicial nos permite achar o valor de $f(t)$ em $t = 0_+$ diretamente da transformada de Laplace de $f(t)$. O teorema do valor inicial não dá o valor de $f(t)$ exatamente no ponto $t = 0$, mas sim o valor da função em um instante um pouco maior que zero.

Teorema do valor inicial. Se $f(t)$ e $df(t)/dt$ são ambos transformáveis segundo Laplace e se $\lim_{s\to\infty} sF(s)$ existe, então

$$f(0+) = \lim_{s\to\infty} sF(s)$$

Para provar este teorema, devemos usar a equação para a transformada $\mathscr{L}_+$ de $df(t)/dt$: ou seja,

$$\mathscr{L}_+\left[\frac{d}{dt}f(t)\right] = sF(s) - f(0-) \tag{2-4}$$

Para um intervalo de tempo $0_+ \le t \le \infty$, quando s tende a infinito e^{-st} tende a zero. (Observe que devemos usar $\mathscr{L}_+$ ao invés de $\mathscr{L}_-$ para esta condição.) Portanto,

$$\lim_{s\to\infty}\int_{0+}^\infty \left[\frac{d}{dt}f(t)\right]e^{-st}\,dt = 0 \tag{2-5}$$

Das equações (2.4) e (2.5), obtemos

$$\lim_{s\to\infty} sF(s) - f(0+) = 0$$

Portanto o teorema está provado.

Ao aplicar o teorema do valor inicial, não existe limitação quanto à localização dos pólos de $sF(s)$. (Portanto, o teorema do valor inicial é válido para a função senoidal.) Freqüentemente os teoremas do valor inicial e do valor final fornecem uma maneira conveniente de conferir a solução. (Observe que estes teoremas nos permitem prever o comportamento do sistema no domínio do tempo sem a necessidade de transformar funções em s de volta para funções no tempo.)

Teorema da integração. A transformada de Laplace da integral de $f(t)$ é dada por

$$\mathscr{L}\left[\int f(t)\,dt\right] = \frac{F(s)}{s} + \frac{f^{-1}(0)}{s} \tag{2-6}$$

onde $f^{-1}(0) = \int f(t)dt$ calculado em $t = 0$.

Observe que se $f(t)$ envolve uma função impulso em $t = 0$, então $f^{-1}(0_+) \neq f^{-1}(0_-)$. Portanto, se $f(t)$ envolve uma função impulso em $t = 0$, então devemos modificar a Eq. (2.6) como segue:

$$\mathscr{L}_+\left[\int f(t)\,dt\right] = \frac{F(s)}{s} + \frac{f^{-1}(0+)}{s}$$

$$\mathscr{L}_-\left[\int f(t)\,dt\right] = \frac{F(s)}{s} + \frac{f^{-1}(0-)}{s}$$

Este teorema pode ser provado como segue: Integrando por partes, nos dá

$$\mathscr{L}\left[\int f(t)\,dt\right] = \int_0^\infty \left[\int f(t)\,dt\right] e^{-st}\,dt$$
$$= \left[\int f(t)\,dt\right]\frac{e^{-st}}{-s}\bigg|_0^\infty - \int_0^\infty f(t)\,\frac{e^{-st}}{-s}\,dt$$
$$= \frac{1}{s}\int f(t)\,dt\bigg|_{t=0} + \frac{1}{s}\int_0^\infty f(t)e^{-st}\,dt$$
$$= \frac{f^{-1}(0)}{s} - \frac{F(s)}{s}$$

Portanto o teorema está provado.

Vemos então que integração no domínio do tempo é convertida em divisão no domínio s. Se o valor inicial da integral é zero, então a transformada de Laplace da integral de $f(t)$ é dada por $F(s)/s$.

Integral de convolução. Considere a transformada de Laplace de

$$\int_0^t f_1(t-\tau)f_2(\tau)\,d\tau$$

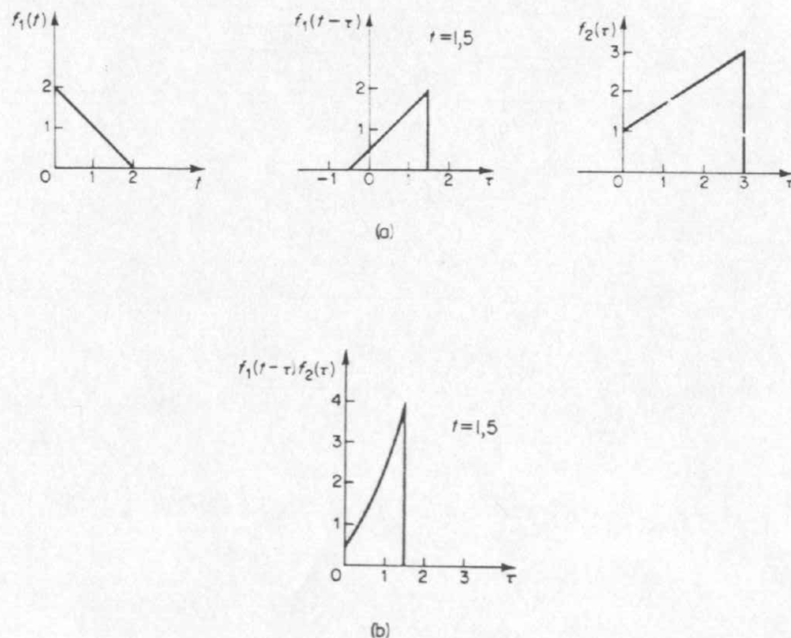


Fig. 2.5 (a) Gráficos de $f_1(t)$, $f_1(t-\tau)$, e $f_2(\tau)$; (b) gráficos de $f_1(t-\tau)f_2(\tau)$.

Esta integral é freqüentemente escrita como

$$f_1(t)*f_2(t)$$

A operação matemática $f_1(t)*f_2(t)$ é chamada *convolução*. Note que, se pusermos $t - \tau = \xi$, então

$$\int_0^t f_1(t-\tau)f_2(\tau)\,d\tau = -\int_t^0 f_1(\xi)f_2(t-\xi)\,d\xi$$
$$= \int_0^t f_1(\tau)f_2(t-\tau)\,d\tau$$

Portanto,

$$f_1(t)*f_2(t) = \int_0^t f_1(t-\tau)f_2(\tau)\,d\tau$$
$$= \int_0^t f_1(\tau)f_2(t-\tau)\,d\tau$$
$$= f_2(t)*f_1(t)$$

A Fig. 2.5(a) mostra curvas de $f_1(t)$, $f_1(t-\tau)$, e $f_2(\tau)$. A Fig. 2.5(b) mostra o produto de $f_1(t-\tau)$ e $f_2(\tau)$. A forma da curva $f_1(t-\tau)f_2(\tau)$ depende de t.

Se $f_1(t)$ e $f_2(t)$ são contínuas por trechos e de ordem exponencial, então a transformada de Laplace de

$$\int_0^t f_1(t-\tau)f_2(\tau)\,d\tau$$

pode ser obtida como segue:

$$\mathscr{L}\left[\int_0^t f_1(t-\tau)f_2(\tau)\,d\tau\right] = F_1(s)F_2(s) \tag{2-7}$$

onde

$$F_1(s) = \int_0^\infty f_1(t)e^{-st}\,dt = \mathscr{L}[f_1(t)]$$
$$F_2(s) = \int_0^\infty f_2(t)e^{-st}\,dt = \mathscr{L}[f_2(t)]$$

Para provar a Eq. (2.7), note que $f_1(t-\tau)1(t-\tau) = 0$ para $\tau > t$. Portanto,

$$\int_0^t f_1(t-\tau)f_2(\tau)\,d\tau = \int_0^\infty f_1(t-\tau)1(t-\tau)f_2(\tau)\,d\tau$$

Então

$$\mathscr{L}\left[\int_0^t f_1(t-\tau)f_2(\tau)\,d\tau\right] = \mathscr{L}\left[\int_0^\infty f_1(t-\tau)1(t-\tau)f_2(\tau)\,d\tau\right]$$
$$= \int_0^\infty e^{-st}\left[\int_0^\infty f_1(t-\tau)1(t-\tau)f_2(\tau)\,d\tau\right]dt$$

	$f(t)$	$F(s)$
1	impulso unitário $\delta(t)$	1
2	degrau unitário $1(t)$	$\frac{1}{s}$
3	t	$\frac{1}{s^2}$
4	e^{-at}	$\frac{1}{s+a}$
5	te^{-at}	$\frac{1}{(s+a)^2}$
6	$\operatorname{sen} \omega t$	$\frac{\omega}{s^2+\omega^2}$
7	$\cos \omega t$	$\frac{s}{s^2+\omega^2}$
8	$t^n \quad (n = 1, 2, 3, \ldots)$	$\frac{n!}{s^{n+1}}$
9	$t^n e^{-at} \quad (n = 1, 2, 3, \ldots)$	$\frac{n!}{(s+a)^{n+1}}$
10	$\frac{1}{b-a}(e^{-at} - e^{-bt})$	$\frac{1}{(s+a)(s+b)}$
11	$\frac{1}{b-a}(be^{-bt} - ae^{-at})$	$\frac{s}{(s+a)(s+b)}$
12	$\frac{1}{ab}\left[1 + \frac{1}{a-b}(be^{-at} - ae^{-bt})\right]$	$\frac{1}{s(s+a)(s+b)}$
13	$e^{-at} \operatorname{sen} \omega t$	$\frac{\omega}{(s+a)^2+\omega^2}$
14	$e^{-at} \cos \omega t$	$\frac{s+a}{(s+a)^2+\omega^2}$
15	$\frac{1}{a^2}(at - 1 + e^{-at})$	$\frac{1}{s^2(s+a)}$
16	$\frac{\omega_n}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \operatorname{sen} \omega_n \sqrt{1-\zeta^2}\, t$	$\frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$
17	$\frac{-1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \operatorname{sen}(\omega_n \sqrt{1-\zeta^2}\, t - \phi)$ $\phi = \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta}$	$\frac{s}{s^2 + 2\zeta\omega_n s + \omega_n^2}$
18	$1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \operatorname{sen}(\omega_n \sqrt{1-\zeta^2}\, t + \phi)$ $\phi = \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta}$	$\frac{\omega_n^2}{s(s^2 + 2\zeta\omega_n s + \omega_n^2)}$

Tabela 2.2 Propriedades das transformadas de Laplace

1	$\mathcal{L}[Af(t)] = AF(s)$
2	$\mathcal{L}[f_1(t) \pm f_2(t)] = F_1(s) \pm F_2(s)$
3	$\mathcal{L}_\pm\left[\frac{d}{dt}f(t)\right] = sF(s) - f(0\pm)$
4	$\mathcal{L}_\pm\left[\frac{d^2}{dt^2}f(t)\right] = s^2F(s) - sf(0\pm) - \dot{f}(0\pm)$
5	$\mathcal{L}_\pm\left[\frac{d^n}{dt^n}f(t)\right] = s^nF(s) - \sum_{k=1}^{n} s^{n-k} \overset{(k-1)}{f}(0\pm)$ onde $\overset{(k-1)}{f}(t) = \frac{d^{k-1}}{dt^{k-1}}f(t)$
6	$\mathcal{L}_\pm\left[\int f(t)\,dt\right] = \frac{F(s)}{s} + \frac{\left[\int f(t)\,dt\right]_{t=0\pm}}{s}$
7	$\mathcal{L}_\pm\left[\iint f(t)\,dt\,dt\right] = \frac{F(s)}{s^2} + \frac{\left[\int f(t)\,dt\right]_{t=0\pm}}{s^2} + \frac{\left[\iint f(t)\,dt\,dt\right]_{t=0\pm}}{s}$
8	$\mathcal{L}_\pm\left[\int \cdots \int f(t)\,(dt)^n\right] = \frac{F(s)}{s^n} + \sum_{k=1}^{n} \frac{1}{s^{n-k+1}}\left[\int \cdots \int f(t)\,(dt)^k\right]_{t=0\pm}$
9	$\mathcal{L}[e^{-at}f(t)] = F(s+a)$
10	$\mathcal{L}[f(t-a)\,1(t-a)] = e^{-as}F(s)$
11	$\mathcal{L}[tf(t)] = -\frac{dF(s)}{ds}$
12	$\mathcal{L}\left[\frac{1}{t}f(t)\right] = \int_s^\infty F(s)\,ds$
13	$\mathcal{L}\left[f\left(\frac{t}{a}\right)\right] = a\,F(as)$

Substituindo $t - \tau = \lambda$ nesta última equação e mudando a ordem de integração, o que é válido neste caso pois $f_1(t)$ e $f_2(t)$ são transformáveis segundo Laplace, obtemos

$$\begin{aligned}
\mathcal{L}\left[\int_0^t f_1(t-\tau)f_2(\tau)\,d\tau\right] &= \int_0^\infty f_1(t-\tau)1(t-\tau)e^{-st}\,dt \int_0^\infty f_2(\tau)d\tau \\
&= \int_0^\infty f_1(\lambda)e^{-s(\lambda+\tau)}\,d\lambda \int_0^\infty f_2(\tau)\,d\tau \\
&= \int_0^\infty f_1(\lambda)e^{-s\lambda}\,d\lambda \int_0^\infty f_2(\tau)e^{-s\tau}\,d\tau \\
&= F_1(s)F_2(s)
\end{aligned}$$

Tabelas de transformadas de Laplace. A Tabela 2.1 fornece uma lista de pares de transformadas de Laplace. A tabela pode ser usada para se obter a transformada de Laplace de uma dada função do tempo ou para obter uma função do tempo correspondendo a uma dada transformada de Laplace. A Tabela 2.1 pode ser adequada para a resolução de problemas comuns no campo de sistemas de controle. A Tabela 2.2 resume uma lista de teoremas e relações de utilidade na teoria da transformada de Laplace.

2.4 TRANSFORMAÇÃO INVERSA DE LAPLACE

O processo matemático de se passar de uma expressão com variável complexa para a expressão no tempo é chamada de *transformação inversa*. A notação para a transformação inversa de Laplace é $\mathscr{L}^{-1}$, de modo que

$$\mathscr{L}^{-1}[F(s)] = f(t)$$

Ao resolver problemas usando o método da transformada de Laplace, nos confrontamos com a pergunta de como determinar $f(t)$ a partir de $F(s)$. Matematicamente, $f(t)$ é determinado de $F(s)$ pela seguinte expressão:

$$f(t) = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} F(s)e^{st}\,ds \qquad (t > 0) \tag{2-8}$$

onde c, a abscissa de convergência, é uma constante real e é escolhida maior que as partes reais de todos os pontos singulares de $F(s)$. Portanto, a trajetória de integração é paralela ao eixo $j\omega$ e está deslocada de um valor c deste. Esta trajetória de integração se situa à direita de todos os pontos singulares.

A integração dada pela Eq. (2.8) parece ser complicada. Por sorte, há métodos mais simples de se achar $f(t)$ a partir de $F(s)$ do que fazendo a integração diretamente. Um método conveniente de se obter transformadas inversas de Laplace é usar uma tabela de transformadas de Laplace. Neste caso, a transformada deve estar em uma forma imediatamente reconhecível em tal tabela. Freqüentemente a função em questão pode não aparecer em tabelas de transformadas de Laplace disponíveis para o engenheiro. Se uma particular transformada $F(s)$ não é achada em um tabela, então podemos expandir $F(s)$ em frações parciais e escrever $F(s)$ em termos de funções simples de s para as quais as transformadas inversas de Laplace são conhecidas.

Note que estes métodos mais simples de achar as transformadas inversas de Laplace são baseados no fato de que a correspondência de uma função temporal e sua transformada inversa de Laplace vale para qualquer função contínua do tempo.*

Método da expansão em frações parciais para determinar transformadas inversas de Laplace. Se $F(s)$, a transformada de Laplace de $f(t)$, é separada em componentes

$$F(s) = F_1(s) + F_2(s) + \cdots + F_n(s)$$

*N. do T.: Fica mais claro se escrevermos ...uma função temporal e a *transformada inversa de Laplace de* $F(s)$ vale para...

e se as transformadas inversas de Laplace de $F_1(s)$, $F_2(s)$, . . . , $F_n(s)$ são conhecidas, então

$$\begin{aligned}\mathscr{L}^{-1}[F(s)] &= \mathscr{L}^{-1}[F_1(s)] + \mathscr{L}^{-1}[F_2(s)] + \cdots + \mathscr{L}^{-1}[F_n(s)] \\ &= f_1(t) + f_2(t) + \cdots + f_n(t)\end{aligned} \tag{2-9}$$

onde $f_1(t)$, $f_2(t)$, . . . , $f_n(t)$ são as transformadas inversas de $F_1(s)$, $F_2(s)$, . . . , $F_n(s)$, respectivamente.

Para problemas em teoria de controle, $F(s)$ é freqüentemente expresso na seguinte forma:

$$F(s) = \frac{B(s)}{A(s)}$$

onde $A(s)$ e $B(s)$ são polinômios em s, e o grau de $B(s)$ não é maior que o de $A(s)$.

Ao aplicar a técnica de expansão em frações parciais para achar a transformada inversa de Laplace de $F(s) = B(s)/A(s)$, devem-se conhecer de antemão as raízes do polinômio do denominador $A(s)$. (Em outras palavras, este método não é aplicável enquanto o polinômio do denominador não foi fatorado.)

A vantagem do método da expansão em frações parciais é que termos individuais de $F(s)$, resultando da expansão na forma de frações parciais, são funções muito simples de s; portanto não necessitamos consultar uma tabela de transformadas de Laplace se memorizarmos vários pares de transformadas de Laplace simples.

Considere $F(s)$ escrito na seguinte forma fatorada:

$$F(s) = \frac{B(s)}{A(s)} = \frac{K(s+z_1)(s+z_2)\cdots(s+z_m)}{(s+p_1)(s+p_2)\cdots(s+p_n)}$$

onde $p_1, p_2, \ldots, p_n$ e $z_1, z_2, \ldots, z_m$ são grandezas reais ou complexas; mas para cada p ou z complexos vai ocorrer o complexo conjugado de p ou z, respectivamente. Aqui supomos que a maior potência de s em $A(s)$ é maior que a de $B(s)$.

Na expansão de $B(s)/A(s)$ em forma de frações parciais, é importante que a maior potência de s em $A(s)$ seja maior que a maior potência de s em $B(s)$. Se este não for o caso, o numerador $B(s)$ deve ser dividido pelo denominador $A(s)$ para resultar um polinômio em s mais um resto (uma razão entre polinômios em s cujo numerador é de grau menor que o denominador). (Para detalhes, veja Exemplo 2.9.)

Expansão em frações parciais quando $F(s)$ tem apenas pólos distintos. Neste caso $F(s)$ pode ser sempre expandido em uma soma de simples frações parciais como segue:

$$F(s) = \frac{B(s)}{A(s)} = \frac{a_1}{s+p_1} + \frac{a_2}{s+p_2} + \cdots + \frac{a_n}{s+p_n} \tag{2-10}$$

onde a_k são constantes. Aqui a_k é chamado de *resíduo* no pólo $s = -p_k$. O valor de a_k pode ser achado multiplicando-se ambos os lados da Eq. (2.10) por $(s + p_k)$ e fazendo $s = -p_k$, resultando

$$\left[\frac{B(s)}{A(s)}(s + p_k)\right]_{s=-p_k} = \left[\frac{a_1}{s + p_1}(s + p_k) + \frac{a_2}{s + p_2}(s + p_k) + \cdots + \frac{a_k}{s + p_k}(s + p_k) + \cdots + \frac{a_n}{s + p_n}(s + p_k)\right]_{s=-p_k} = a_k$$

Vemos que todos os termos expandidos são cancelados, com exceção de a_k. Portanto, o resíduo a_k é achado de

$$a_k = \left[\frac{B(s)}{A(s)}(s + p_k)\right]_{s=-p_k} \tag{2-11}$$

Note que desde que $f(t)$ é uma função real do tempo, se p_1 e p_2 são complexos conjugados, então os resíduos a_1 e a_2 são também complexos conjugados. Apenas um dos valores conjugados a_1 ou a_2 necessita ser calculado uma vez que o outro é conhecido automaticamente.

Com referência à Eq. (2.9) e percebendo que

$$\mathscr{L}^{-1}\left[\frac{a_k}{s + p_k}\right] = a_k e^{-p_k t}$$

obtemos $f(t) = \mathscr{L}^{-1}[F(s)]$ como segue:

$$f(t) = a_1 e^{-p_1 t} + a_2 e^{-p_2 t} + \cdots + a_n e^{-p_n t} \qquad (t \geq 0)$$

Exemplo 2.8 Determine a transformada inversa de Laplace de

$$F(s) = \frac{s + 3}{(s + 1)(s + 2)}$$

A expansão em frações parciais de $F(s)$ é

$$F(s) = \frac{s + 3}{(s + 1)(s + 2)} = \frac{a_1}{s + 1} + \frac{a_2}{s + 2}$$

onde a_1 e a_2 são achados usando-se a Eq. (2.11) como segue:

$$a_1 = \left[\frac{s + 3}{(s + 1)(s + 2)}(s + 1)\right]_{s=-1} = 2$$

$$a_2 = \left[\frac{s + 3}{(s + 1)(s + 2)}(s + 2)\right]_{s=-2} = -1$$

Portanto,

$$\begin{aligned} f(t) &= \mathscr{L}^{-1}[F(s)] \\ &= \mathscr{L}^{-1}\left[\frac{2}{s + 1}\right] + \mathscr{L}^{-1}\left[\frac{-1}{s + 2}\right] \\ &= 2e^{-t} - e^{-2t} \qquad (t \geq 0) \end{aligned}$$

Exemplo 2.9 Determine a transformada de Laplace inversa de

$$G(s) = \frac{s^3 + 5s^2 + 9s + 7}{(s + 1)(s + 2)}$$

Dividindo o numerador pelo denominador, obtemos

$$G(s) = s + 2 + \frac{s + 3}{(s + 1)(s + 2)}$$

O terceiro termo do lado direito desta última equação é $F(s)$ no Exemplo 2.8. Observe que a transformada de Laplace da função impulso unitário $\delta(t)$ é 1 e a transformada de Laplace de $d\delta(t)/dt$ é s. Obtemos então a transformada inversa de Laplace de $G(s)$ como segue:

$$g(t) = \frac{d}{dt}\delta(t) + 2\,\delta(t) + 2e^{-t} - e^{-2t} \qquad (t \geq 0-)$$

Expansão em frações parciais quando $F(s)$ apresenta pólos complexos conjugados. Se p_1 e p_2 são pólos complexos conjugados, então a seguinte expansão pode ser usada:

$$F(s) = \frac{B(s)}{A(s)} = \frac{\alpha_1 s + \alpha_2}{(s + p_1)(s + p_2)} + \frac{a_3}{s + p_3} + \cdots + \frac{a_n}{s + p_n} \tag{2-12}$$

Os valores de α_1 e α_2 são determinados multiplicando-se ambos os lados da Eq. (2.12) por $(s + p_1)(s + p_2)$ e fazendo $s = -p_1$, dando

$$\begin{aligned} &\left[\frac{B(s)}{A(s)}(s + p_1)(s + p_2)\right]_{s=-p_1} \\ &\quad = \left[(\alpha_1 s + \alpha_2) + \frac{a_3}{s + p_3}(s + p_1)(s + p_2) + \cdots + \frac{a_n}{s + p_n}(s + p_1)(s + p_2)\right]_{s=-p_1} \end{aligned}$$

Vemos que todos os termos expandidos são cancelados com exceção do termo $(\alpha_1 s + \alpha_2)$. Portanto,

$$(\alpha_1 s + \alpha_2)_{s=-p_1} = \left[\frac{B(s)}{A(s)}(s + p_1)(s + p_2)\right]_{s=-p_1} \tag{2-13}$$

Como p_1 é uma grandeza complexa, ambos os lados da Eq. (2.13) são grandezas complexas. Igualando as partes reais de ambos os lados da Eq. (2.13), obtemos uma equação. Da mesma forma, igualando as partes imaginárias de ambos os lados da Eq. (2.13), obtemos uma outra equação. Dessas duas equações é possível determinar α_1 e α_2.

Exemplo 2.10 Determine a transformada inversa de Laplace de

$$F(s) = \frac{s + 1}{s(s^2 + s + 1)}$$

$F(s)$ pode ser expandida da seguinte forma:

$$\frac{s+1}{s(s^2+s+1)} = \frac{\alpha_1 s + \alpha_2}{s^2+s+1} + \frac{a}{s} \tag{2-14}$$

Para determinar α_1 e α_2, note que

$$s^2 + s + 1 = (s + 0{,}5 + j0{,}866)(s + 0{,}5 - j0{,}866)$$

Multiplicando ambos os lados da Eq. (2.14) por $(s^2 + s + 1)$ e impondo $s = -0{,}5 - j0{,}866$, obtemos

$$\left(\frac{s+1}{s}\right)_{s=-0,5-j0,866} = (\alpha_1 s + \alpha_2)_{s=-0,5-j0,866}$$

ou

$$\frac{0{,}5 - j0{,}866}{-0{,}5 - j0{,}866} = \alpha_1(-0{,}5 - j0{,}866) + \alpha_2$$

que pode ser simplificado como visto a seguir:

$$0{,}5 - j0{,}866 = \alpha_1(0{,}25 + j0{,}866 - 0{,}75) + \alpha_2(-0{,}5 - j0{,}866)$$

Igualando as partes reais e imaginárias de ambos os lados desta equação, respectivamente, obtemos

$$-0{,}5\alpha_1 - 0{,}5\alpha_2 = 0{,}5$$

$$0{,}866\alpha_1 - 0{,}866\alpha_2 = -0{,}866$$

ou

$$\alpha_1 + \alpha_2 = -1$$

$$\alpha_1 - \alpha_2 = -1$$

Resultando

$$\alpha_1 = -1, \qquad \alpha_2 = 0$$

Para determinar a, multiplicamos ambos os lados da Eq. (2.14) por s e fazemos $s = 0$, resultando

$$a = \left[\frac{s(s+1)}{s(s^2+s+1)}\right]_{s=0} = 1$$

Portanto,

$$F(s) = \frac{-s}{s^2+s+1} + \frac{1}{s}$$

$$= \frac{1}{s} - \frac{s+0{,}5}{(s+0{,}5)^2 + 0{,}866^2} + \frac{0{,}5}{(s+0{,}5)^2 + 0{,}866^2}$$

A transformada inversa de Laplace de $F(s)$ é então dada por

$$f(t) = \mathcal{L}^{-1}[F(s)]$$

$$= 1 - e^{-0,5t}\cos 0{,}866t + 0{,}578e^{-0,5t}\,\mathrm{sen}\,0{,}866t \qquad (t \geq 0)$$

Expansão em frações parciais quando $F(s)$ apresenta pólos múltiplos. Considere $F(s) = B(s)/A(s)$, onde $A(s) = 0$ tem raízes p_1 de multiplicidade r. (As outras raízes são supostas distintas.) $A(s)$ pode então ser escrita como

$$A(s) = (s + p_1)^r (s + p_{r+1})(s + p_{r+2}) \cdots (s + p_n)$$

A expansão em frações parciais de $F(s)$ é

$$F(s) = \frac{B(s)}{A(s)} = \frac{b_r}{(s+p_1)^r} + \frac{b_{r-1}}{(s+p_1)^{r-1}} + \cdots + \frac{b_1}{s+p_1}$$
$$+ \frac{a_{r+1}}{s+p_{r+1}} + \frac{a_{r+2}}{s+p_{r+2}} + \cdots + \frac{a_n}{s+p_n} \tag{2-15}$$

onde $b_r, b_{r-1}, \ldots, b_1$ são dados por

$$b_r = \left[\frac{B(s)}{A(s)}(s+p_1)^r\right]_{s=-p_1}$$

$$b_{r-1} = \left\{\frac{d}{ds}\left[\frac{B(s)}{A(s)}(s+p_1)^r\right]\right\}_{s=-p_1}$$

$$\vdots$$

$$b_{r-j} = \frac{1}{j!}\left\{\frac{d^j}{ds^j}\left[\frac{B(s)}{A(s)}(s+p_1)^r\right]\right\}_{s=-p_1}$$

$$\vdots$$

$$b_1 = \frac{1}{(r-1)!}\left\{\frac{d^{r-1}}{ds^{r-1}}\left[\frac{B(s)}{A(s)}(s+p_1)^r\right]\right\}_{s=-p_1}$$

Estas relações para os b's podem ser obtidas da seguinte forma: Multiplicando ambos os lados da Eq. (2.15) por $(s + p_1)^r$ e fazendo s tender a $-p_1$, temos

$$b_r = \left[\frac{B(s)}{A(s)}(s+p_1)^r\right]_{s=-p_1}$$

Se multiplicarmos ambos os lados da Eq. (2.15) por $(s + p_1)^r$ e então derivarmos com relação a s,

$$\frac{d}{ds}\left[\frac{B(s)}{A(s)}(s+p_1)^r\right] = b_r\frac{d}{ds}\left[\frac{(s+p_1)^r}{(s+p_1)^r}\right] + b_{r-1}\frac{d}{ds}\left[\frac{(s+p_1)^r}{(s+p_1)^{r-1}}\right]$$
$$+ \cdots + b_1\frac{d}{ds}\left[\frac{(s+p_1)^r}{s+p_1}\right] + a_{r+1}\frac{d}{ds}\left[\frac{(s+p_1)^r}{s+p_{r+1}}\right]$$
$$+ \cdots + a_n\frac{d}{ds}\left[\frac{(s+p_1)^r}{s+p_n}\right]$$

O primeiro termo do lado direito desta última equação é igual a zero. O segundo termo é igual a b_{r-1}. Cada um dos outros termos contém alguma potência de $(s + p_1)$ como fator, resultando que quando s tende ao valor $-p_1$, estes termos se anulam. Portanto,

$$b_{r-1} = \lim_{s\to -p_1}\frac{d}{ds}\left[\frac{B(s)}{A(s)}(s+p_1)^r\right]$$
$$= \left\{\frac{d}{ds}\left[\frac{B(s)}{A(s)}(s+p_1)^r\right]\right\}_{s=-p_1}$$

Da mesma forma, fazendo sucessivas diferenciações com relação a s e fazendo s tender a $-p_1$, obtemos equações para os b_{r-j}.

Note que a transformada inversa de Laplace de $1/(s+p_1)^n$ é dada por

$$\mathscr{L}^{-1}\left[\frac{1}{(s+p_1)^n}\right] = \frac{t^{n-1}}{(n-1)!}e^{-p_1 t}$$

As constantes $a_{r+1}, a_{r+2}, \ldots, a_n$ na Eq. (2.15) são determinadas a partir de

$$a_k = \left[\frac{B(s)}{A(s)}(s+p_k)\right]_{s=-p_k} \qquad (k = r+1, r+2, \ldots, n)$$

A transformada inversa de Laplace de $F(s)$ é então obtida como visto a seguir:

$$f(t) = \mathscr{L}^{-1}[F(s)] = \left[\frac{b_r}{(r-1)!}t^{r-1} + \frac{b_{r-1}}{(r-2)!}t^{r-2} + \cdots + b_2 t + b_1\right]e^{-p_1 t}$$
$$+ a_{r+1}e^{-p_{r+1}t} + a_{r+2}e^{-p_{r+2}t} + \cdots + a_n e^{-p_n t} \qquad (t \geq 0)$$

Exemplo 2.11 Determine a transformada inversa de Laplace da seguinte função $F(s)$:

$$F(s) = \frac{s^2+2s+3}{(s+1)^3}$$

Expandindo $F(s)$ em frações parciais, obtemos

$$F(s) = \frac{B(s)}{A(s)} = \frac{b_3}{(s+1)^3} + \frac{b_2}{(s+1)^2} + \frac{b_1}{s+1}$$

onde b_3, b_2 e b_1 são determinados como visto a seguir:

$$b_3 = \left[\frac{B(s)}{A(s)}(s+1)^3\right]_{s=-1}$$
$$= (s^2+2s+3)_{s=-1} = 2$$
$$b_2 = \left\{\frac{d}{ds}\left[\frac{B(s)}{A(s)}(s+1)^3\right]\right\}_{s=-1}$$
$$= \left[\frac{d}{ds}(s^2+2s+3)\right]_{s=-1}$$
$$= (2s+2)_{s=-1} = 0$$
$$b_1 = \frac{1}{(3-1)!}\left\{\frac{d^2}{ds^2}\left[\frac{B(s)}{A(s)}(s+1)^3\right]\right\}_{s=-1}$$
$$= \frac{1}{2!}\left[\frac{d^2}{ds^2}(s^2+2s+3)\right]_{s=-1}$$
$$= \frac{1}{2}(2) = 1$$

Portanto, obtemos

$$f(t) = \mathscr{L}^{-1}[F(s)]$$
$$= \mathscr{L}^{-1}\left[\frac{2}{(s+1)^3}\right] + \mathscr{L}^{-1}\left[\frac{1}{s+1}\right]$$
$$= (t^2+1)e^{-t} \qquad (t \geq 0)$$

2.5 SOLUÇÃO DE EQUAÇÕES DIFERENCIAIS LINEARES PELO MÉTODO DA TRANSFORMADA DE LAPLACE

Nas Seções 2.1 a 2.4 apresentamos alguns conceitos e técnicas do método da transformada de Laplace. Esta seção apresenta o uso do método da transformada de Laplace na resolução de equações diferenciais lineares.

O método da transformada de Laplace fornece a solução completa (a solução particular mais a solução complementar) de equações diferenciais lineares. Métodos clássicos para achar a solução completa de uma equação diferencial requerem a determinação das constantes de integração pelo uso das condições iniciais. No caso do método da transformada de Laplace, a determinação das constantes de integração a partir das condições iniciais não é necessária uma vez que as condições iniciais são automaticamente incluídas na transformada de Laplace da equação diferencial.

Se todas as condições iniciais são nulas, então a transformada de Laplace da equação diferencial é obtida simplesmente substituindo-se d/dt por s, d^2/dt^2 por s^2 etc.

Ao resolver equações diferenciais lineares pelo método da transformada de Laplace, procedemos de acordo com as duas etapas descritas a seguir:

1. Tomando a transformada de Laplace de cada termo na equação diferencial linear dada, converte-se a equação diferencial em uma equação algébrica em s e se obtém a expressão para a transformada de Laplace da variável dependente através de um rearranjo da equação algébrica.
2. A solução temporal da equação diferencial é obtida achando-se a transformada inversa de Laplace da variável dependente.

Considere como um exemplo a seguinte equação diferencial:

$$m\ddot{x} + kx = f(t) \tag{2-16}$$

onde $f(t)$ representa a função de excitação ou entrada. Tomando a transformada de Laplace de cada termo na Eq. (2.16), obtemos

$$\mathscr{L}[m\ddot{x}] = m[s^2X(s) - sx(0) - \dot{x}(0)]$$
$$\mathscr{L}[kx] = kX(s)$$
$$\mathscr{L}[f(t)] = F(s)$$

A transformada de Laplace da Eq. (2.16) pode então ser escrita como

$$(ms^2 + k)X(s) - msx(0) - m\dot{x}(0) = F(s) \tag{2-17}$$

Resolvendo a Eq. (2.17) para obter $X(s)$, obtemos

$$X(s) = \frac{F(s)}{ms^2 + k} + \frac{msx(0) + m\dot{x}(0)}{ms^2 + k} \tag{2-18}$$

O primeiro termo do lado direito da Eq. (2.18) representa a solução da equação diferencial quando as condições iniciais são todas nulas (solução particular). O segundo termo do lado direito desta equação representa o efeito das condições iniciais (solução complementar). A solução no tempo da equação diferencial é obtida tomando-se a transformada inversa de Laplace de $X(s)$ como visto a seguir:

$$x(t) = \mathscr{L}^{-1}\left[\frac{F(s)}{ms^2 + k}\right] + \mathscr{L}^{-1}\left[\frac{msx(0) + m\dot{x}(0)}{ms^2 + k}\right] \tag{2-19}$$

Como um exemplo, se $f(t)$ é uma função degrau unitário, então $F(s) = 1/s$ e a Eq. (2.19) se torna

$$x(t) = \mathscr{L}^{-1}\left[\frac{1}{s(ms^2 + k)}\right] + \mathscr{L}^{-1}\left[\frac{msx(0) + m\dot{x}(0)}{ms^2 + k}\right]$$
$$= \left(\frac{1}{k} - \frac{1}{k}\cos\sqrt{\frac{k}{m}}t\right) + \left[x(0)\cos\sqrt{\frac{k}{m}}t + \dot{x}(0)\sqrt{\frac{m}{k}}\operatorname{sen}\sqrt{\frac{k}{m}}t\right]$$

Pode-se ver que as condições iniciais $x(0)$ e $\dot{x}(0)$ aparecem na solução. Portanto, $x(t)$ não tem constantes indeterminadas.

PROBLEMAS ILUSTRATIVOS E SOLUÇÕES

Problema A.2.1 Determine os pólos da seguinte função $F(s)$:

$$F(s) = \frac{1}{1 - e^{-s}}$$

Solução. Os pólos são achados a partir de

$$e^{-s} = 1$$

ou

$$e^{-(\sigma + j\omega)} = e^{-\sigma}(\cos\omega - j\operatorname{sen}\omega) = 1$$

De onde segue-se que $\sigma = 0$, $\omega = \pm 2n\pi$ $(n = 0, 1, 2, \ldots)$. Portanto, os pólos estão localizados em

$$s = \pm j2n\pi \qquad (n = 0, 1, 2, \ldots)$$

Problema A.2.2 A derivada da função impulso unitário $\delta(t)$ é chamada uma função *doublet* unitária. (Portanto, a integral da função *doublet* unitária é a função impulso unitário.) Matematicamente, um exemplo da função *doublet* unitário $u_2(t)$ pode ser dado por

$$u_2(t) = \lim_{t_0 \to 0} \frac{1(t) - 2[1(t - t_0)] + 1(t - 2t_0)}{t_0^2}$$

Obtenha a transformada de Laplace de $u_2(t)$.

Solução. A transformada de Laplace de $u_2(t)$ é dada por

$$\mathscr{L}[u_2(t)] = \lim_{t_0 \to 0} \frac{1}{t_0^2}\left(\frac{1}{s} - \frac{2}{s}e^{-t_0 s} + \frac{1}{s}e^{-2t_0 s}\right)$$
$$= \lim_{t_0 \to 0} \frac{1}{t_0^2 s}\left[1 - 2\left(1 - t_0 s + \frac{t_0^2 s^2}{2} + \cdots\right) + \left(1 - 2t_0 s + \frac{4t_0^2 s^2}{2} + \cdots\right)\right]$$
$$= \lim_{t_0 \to 0} \frac{1}{t_0^2 s}\left[t_0^2 s^2 + (\text{termos de ordem superior em } t_0 s)\right] = s$$

Problema A.2.3 O método gráfico de determinar resíduos poupa tempo principalmente nos casos em que o sistema tem muitos pólos e zeros complexos conjugados.

Considere a seguinte função $F(s)$:

$$F(s) = \frac{s + z}{s(s + p_1)(s + p_2)}$$

onde p_1 e p_2 são grandezas complexas conjugadas e z é uma grandeza real. A função $F(s)$ tem três pólos, localizados em $s = 0$, $s = -p_1$, e $s = -p_2$, e um zero finito em $s = -z$. Se estes pólos e zero são postos no gráfico do plano complexo, então os resíduos na expansão em frações parciais podem ser calculados graficamente. A expansão em frações parciais de $F(s)$ é

$$F(s) = \frac{a_1}{s} + \frac{a_2}{s + p_1} + \frac{a_3}{s + p_2}$$

onde $a_1 = z/(p_1 p_2)$ e a_2 é dado por

$$a_2 = \left[\frac{s + z}{s(s + p_2)}\right]_{s=-p_1} = \frac{z - p_1}{(-p_1)(p_2 - p_1)}$$

Aqui $(z - p_1)$ pode ser considerado um vetor que vai de $-z$ a $-p_1$. Com referência à Fig. 2.6, vemos que

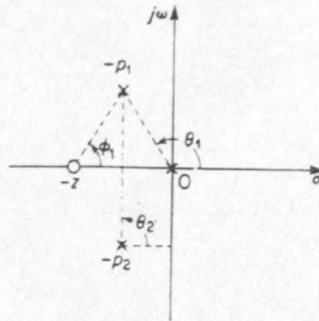


Fig. 2.6 Gráfico de pólos e zeros.

$$z - p_1 = |z - p_1| \underline{/\phi_1}$$

Da mesma forma,

$$-p_1 = |0 - p_1| \underline{/\theta_1}$$

e

$$p_2 - p_1 = |p_2 - p_1| \underline{/\theta_2}$$

Os módulos $|z - p_1|$, $|0 - p_1|$ e $|p_2 - p_1|$ são medidos diretamente. Os ângulos ϕ_1, θ_1 e θ_2 são determinados com um compasso. Então, o resíduo a_2 é determinado como sendo

$$a_2 = \frac{|z - p_1|}{|0 - p_1| \cdot |p_2 - p_1|} \underline{/\phi_1 - \theta_1 - \theta_2}$$

Como a_3 é o complexo conjugado de a_2, segue-se que

$$a_3 = \frac{|z - p_1|}{|0 - p_1| \cdot |p_2 - p_1|} \underline{/-\phi_1 + \theta_1 + \theta_2}$$

(Observe que o resíduo no pólo na origem pode ser achado facilmente através de meios algébricos.)

Determine os resíduos nos pólos quando $z = 1$, $p_1 = 0{,}5 + j0{,}866$, e $p_2 = 0{,}5 - j0{,}866$.

Solução. O resíduo a_1 é

$$a_1 = \left[\frac{s + z}{(s + p_1)(s + p_2)}\right]_{s=0} = \frac{z}{p_1 p_2} = 1$$

Graficamente $(z - p_1)$, $(0 - p_1)$, e $(p_2 - p_1)$ são

$$z - p_1 = 1 \underline{/60^\circ}$$
$$0 - p_1 = 1 \underline{/120^\circ}$$
$$p_2 - p_1 = 1{,}73 \underline{/90^\circ}$$

Portanto,

$$a_2 = \frac{1}{1 \times 1{,}73} \underline{/60^\circ - 120^\circ - 90^\circ}$$
$$= 0{,}578 \underline{/-150^\circ}$$
$$= 0{,}578(\cos 150^\circ - j \operatorname{sen} 150^\circ)$$
$$= -0{,}5 - j0{,}289$$

Portanto, a_3, que é o complexo conjugado de a_2, é obtido como

$$a_3 = -0{,}5 + j0{,}289$$

Este método gráfico de calcular os resíduos é muito conveniente quando o engenheiro deseja uma estimativa a grosso modo dos valores do resíduo, sem entrar em cálculos laboriosos. Observe, entretanto, que este cálculo gráfico não se aplica a pólos múltiplos.

Problema A.2.4 Determine a transformada de Laplace da seguinte equação diferencial:

$$\ddot{x} + 3\dot{x} + 6x = 0, \qquad x(0) = 0, \qquad \dot{x}(0) = 3$$

Achando a transformada inversa de Laplace de $X(s)$, obtenha a solução no tempo $x(t)$.

Solução. A transformada de Laplace da equação diferencial é

$$s^2 X(s) - sx(0) - \dot{x}(0) + 3sX(s) - 3x(0) + 6X(s) = 0$$

Substituindo as condições iniciais e resolvendo para $X(s)$,

$$X(s) = \frac{3}{s^2 + 3s + 6} = \frac{2\sqrt{3}}{\sqrt{5}} \frac{\frac{\sqrt{15}}{2}}{(s + 1{,}5)^2 + \left(\frac{\sqrt{15}}{2}\right)^2}$$

A transformada inversa de Laplace de $X(s)$ é

$$x(t) = \frac{2\sqrt{3}}{\sqrt{5}} e^{-1{,}5t} \operatorname{sen}\left(\frac{\sqrt{15}}{2} t\right)$$

Problema A.2.5 Considere o sistema mecânico visto na Fig. 2.7. Suponha que o sistema é posto em movimento por uma força impulso unitário. Determine a oscilação resultante. Suponha que o sistema está inicialmente em repouso.

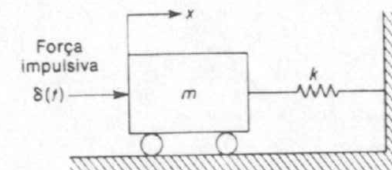


Fig. 2.7 Sistema mecânico.

Solução. O sistema é excitado por uma entrada impulsiva. Portanto.

$$m\ddot{x} + kx = \delta(t)$$

Tomando a transformada de Laplace em ambos os lados desta equação, obtemos

$$m[s^2X(s) - sx(0) - \dot{x}(0)] + kX(s) = 1$$

Substituindo as condições iniciais $x(0) = 0$ e $\dot{x}(0) = 0$ nesta última equação e determinando $X(s)$, obtemos

$$X(s) = \frac{1}{ms^2 + k}$$

A transformada inversa de Laplace de $X(s)$ se torna

$$x(t) = \frac{1}{\sqrt{mk}} \operatorname{sen} \sqrt{\frac{k}{m}}\, t$$

A oscilação é um movimento harmônico simples. A amplitude da oscilação é $1/\sqrt{mk}$.

Problema A.2.6 Considere uma função periódica $f_p(t)$ com período T. Se $f(t)$ é uma função do tempo tal que $f(t) = f_p(t)$ no intervalo $0 < t < T$ e $f(t) = 0$ fora do intervalo $0 < t < T$, então $f_p(t)$ pode ser expressa como

$$f_p(t) = f(t) + f(t - T) + f(t - 2T) + f(t - 3T) + \cdots$$

Determine a transformada de Laplace de $f_p(t)$.

Solução. Defina a transformada de Laplace de $f(t)$ por $F(s)$ ou

$$\mathscr{L}[f(t)] = F(s)$$

Então a transformada de Laplace de $f_p(t)$ é obtida como sendo

$$\begin{aligned}\mathscr{L}[f_p(t)] = F_p(s) &= F(s) + e^{-Ts}F(s) + e^{-2Ts}F(s) + e^{-3Ts}F(s) + \cdots \\ &= (1 + e^{-Ts} + e^{-2Ts} + e^{-3Ts} + \cdots)F(s) \\ &= \frac{F(s)}{1 - e^{-Ts}}\end{aligned}$$

Portanto, a transformada de Laplace de $f_p(t)$ pode ser obtida multiplicando-se $F(s)$ por $1/(1 - e^{-Ts})$.

PROBLEMAS

Problema B.2.1 Determine a transformada de Laplace da função $f(t)$ definida por

$$\begin{aligned} f(t) &= \frac{1}{a^2} && \text{para } 0 < t < a \\ &= -\frac{1}{a^2} && \text{para } a < t < 2a \\ &= 0 && \text{para } t < 0,\ 2a < t \end{aligned}$$

Determine também o valor limite de $F(s)$ quando a tende a zero.

Problema B.2.2 Determine as transformadas de Laplace das seguintes funções. Suponha que $f(t) = 0$ para $t < 0$.

1. $f(t) = 0{,}03(1 - \cos 2t)$
2. $f(t) = e^{-0{,}4t} \cos 12t$
3. $f(t) = \operatorname{sen}\left(5t + \dfrac{\pi}{3}\right)$
4. $f(t) = t^n$
5. $f(t) = t^n e^{at}$

Problema B.2.3 Determine as transformadas inversas de Laplace das seguintes funções:

1. $F(s) = \dfrac{1}{s(s + 1)}$
2. $F(s) = \dfrac{s + 1}{(s + 2)(s + 3)}$
3. $F(s) = \dfrac{e^{-s}}{s - 1}$
4. $F(s) = \dfrac{(s + 3)(s + 4)(s + 5)}{(s + 1)(s + 2)}$
5. $F(s) = \dfrac{s}{(s + 1)^2(s + 2)}$

Problema B.2.4 Usando o teorema do valor final, ache o valor final de $f(t)$ cuja transformada de Laplace é dada por

$$F(s) = \frac{10}{s(s + 1)}$$

Verifique este resultado tomando a transformada inversa de Laplace de $F(s)$ e fazendo $t \to \infty$.

Problema B.2.5 Dado

$$F(s) = \frac{1}{(s + 2)^2}$$

determine os valores de $f(0_+)$ e $\dot{f}(0_+)$. (Use o teorema do valor inicial.)

Problema B.2.6 Resolva a seguinte equação diferencial:

$$a\dot{x} + bx = k, \qquad x(0) = x_0$$

onde a, b e k são constantes.

Problema B.2.7 Resolva a seguinte equação diferencial:

$$2\ddot{x} + 7\dot{x} - 3x = 0, \qquad x(0) = x_0, \qquad \dot{x}(0) = 0$$

Problema B.2.8 Resolva a seguinte equação diferencial:

$$\ddot{x} + 2\zeta\omega_n\dot{x} + \omega_n^2 x = 0, \qquad x(0) = a, \qquad \dot{x}(0) = b$$

onde a e b são constantes.

Problema B.2.9 Considere o sistema visto na Fig. 2.7. O sistema está inicialmente em repouso. Suponha que o carro é posto em movimento por uma força impulsiva cuja magnitude (área) é unitária. É possível parar o carro usando-se uma outra destas forças impulsivas?

3

Base Matemática — Matrizes

3.1 INTRODUÇÃO

Na dedução de modelos matemáticos referentes a sistemas de controle modernos, verificamos que as respectivas equações diferenciais podem-se tornar muito complicadas devido à multiplicidade de entradas e saídas. De fato, o número de entradas e saídas de um sistema complexo pode-se elevar a centenas. Para simplificar as expressões matemáticas das equações do sistema, é vantajoso usar a notação vetor-matriz. Em trabalhos teóricos, a simplicidade de notação ganha pelo uso de operações com vetores-matrizes é mais conveniente e, de fato, essencial para a análise e a síntese de sistemas de controle moderno.

A utilização da notação vetor-matriz nos possibilita tratar com problemas grandes e complexos com facilidade, seguindo o formato sistemático de representar as equações do sistema e calculá-las com o uso de computadores.

O principal objetivo deste capítulo é apresentar definições de matrizes e a álgebra matricial básica necessária para a análise subseqüente de sistemas de controle.

3.2 DEFINIÇÕES DE MATRIZES

Esta seção apresenta definições de matrizes que serão usadas em todo o resto do livro.

Matriz. Uma matriz é definida como um arranjo retangular de elementos, os quais podem ser números reais, números complexos, funções ou operadores. O número de colunas, em geral, não é necessariamente igual ao número de linhas. Considere a segunda matriz:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

onde a_{ij} indica o (i, j)ésimo elemento da matriz **A**. Esta matriz possui n linhas e m colunas e é denominada uma matriz $n \times m$. O primeiro índice representa o número da linha e o segundo índice o número da coluna. A matriz **A** é muitas vezes escrita (a_{ij}).

Igualdade de duas matrizes. Duas matrizes são ditas iguais se e somente se os seus elementos correspondentes são iguais. Note que matrizes iguais devem possuir o mesmo número de linhas e o mesmo número de colunas.

Vetor. Uma matriz possuindo apenas uma coluna tal como

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

é denominada um vetor *coluna*. Um vetor coluna com n elementos é chamado vetor n-dimensional ou, simplesmente, vetor n.

Uma matriz possuindo apenas uma linha tal como

$$[x_1 \quad x_2 \quad \cdots \quad x_n]$$

é chamada um vetor *linha*.

Matriz quadrada. Uma matriz quadrada é uma matriz na qual o número de linhas é igual ao número de colunas. A matriz quadrada é muitas vezes denominada *matriz de ordem n*, onde n é o número de linhas (ou colunas).

Matriz diagonal. Se todos os elementos de uma matriz quadrada **A**, excetuando-se apenas os elementos da diagonal principal, são nulos, **A** é denominada matriz *diagonal* e é escrita como

$$\mathbf{A} = \begin{bmatrix} a_{11} & & & & 0 \\ & a_{22} & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & a_{nn} \end{bmatrix} = (a_{ij}\,\delta_{ij})$$

onde os δ_{ij} são os deltas de Kronecker definidos por

$$\begin{aligned} \delta_{ij} &= 1 \quad \text{se } i = j \\ &= 0 \quad \text{se } i \neq j \end{aligned}$$

Note que todos os elementos não explicitamente escritos na matriz precedente são nulos. A matriz diagonal muitas vezes é escrita

$$\operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn})$$

Matriz identidade ou matriz unidade. A matriz identidade ou matriz unidade **I** é uma matriz cujos elementos da diagonal principal são iguais a um e todos os outros elementos são iguais a zero; isto é,

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \operatorname{diag}(1, 1, \ldots, 1)$$

Matriz zero (ou nula). Uma matriz zero é uma matriz na qual todos os elementos são zeros.

Determinante de uma matriz. Para cada matriz quadrada, existe um determinante. O determinante possui as seguintes propriedades:

1. Se duas linhas ou colunas quaisquer consecutivas são permutadas, o determinante muda de sinal.
2. Se qualquer linha ou coluna possuir apenas elementos zeros, então o valor do determinante é zero.
3. Se os elementos de qualquer linha (ou qualquer coluna) são exatamente k vezes aqueles de uma outra linha (ou uma outra coluna), então o valor do determinante é zero.
4. Se a qualquer linha (ou qualquer coluna) qualquer constante vezes uma outra linha (ou coluna) é adicionada, o valor do determinante permanece inalterado.
5. Se um determinante é multiplicado por uma constante, então apenas uma linha (ou uma coluna) é multiplicada pela constante. Note, entretanto, que o determinante de k vezes uma matriz **A** $n \times n$ é k^n vezes o determinante de **A** ou

$$|k\mathbf{A}| = k^n |\mathbf{A}|$$

6. O determinante do produto de duas matrizes quadradas **A** e **B** é o produto dos determinantes ou

$$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$$

Matriz singular. Uma matriz quadrada é denominada *singular* se o determinante associado é zero. Em uma matriz singular nem todas as linhas (ou nem todas as colunas) são independentes de cada outra.

Matriz não singular. Uma matriz quadrada é denominada *não singular* se o determinante associado é diferente de zero.

Transposta. Se as linhas e colunas de uma matriz **A** $n \times m$ são intertrocadas, a matriz $m \times n$ resultante é denominada *transposta* da matriz **A**. A transposta da matriz **A** é denotada por **A**′. Se **A** é dada por

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

então **A**′ é dada por

$$\mathbf{A}' = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{bmatrix}$$

Note que $(\mathbf{A}')' = \mathbf{A}$.

Matriz simétrica. Se uma matriz quadrada **A** é igual à sua transposta, ou seja,

$$\mathbf{A} = \mathbf{A}'$$

então a matriz **A** é denominada matriz *simétrica*.

Matriz anti-simétrica. Se uma matriz quadrada **A** é igual ao negativo de sua transposta, ou seja,

$$\mathbf{A} = -\mathbf{A}'$$

então a matriz **A** é denominada matriz *anti-simétrica*.

Matriz conjugada. Se os elementos complexos de uma matriz **A** são substituídos pelos seus respectivos conjugados, então a matriz resultante é chamada a *conjugada* de **A**. A conjugada de **A** é denotada por $\bar{\mathbf{A}} = (\bar{a}_{ij})$, onde $\bar{a}_{ij}$ é o complexo conjugado de a_{ij}. Por exemplo, se **A** é dada por

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ -1+j & -3-j3 & -1+j4 \\ -1+j & -1 & -2+j3 \end{bmatrix}$$

então

$$\bar{\mathbf{A}} = \begin{bmatrix} 0 & 1 & 0 \\ -1-j & -3+j3 & -1-j4 \\ -1-j & -1 & -2-j3 \end{bmatrix}$$

Transposta conjugada. A transposta conjugada é a conjugada da transposta de uma matriz. Dada uma matriz **A**, a transposta conjugada é denotada por $\bar{\mathbf{A}}'$ ou $\mathbf{A}^*$; isto é,

$$\bar{\mathbf{A}}' = \mathbf{A}^* = (\bar{a}_{ji})$$

Por exemplo, se **A** é dada por

$$\mathbf{A} = \begin{bmatrix} 1 & j2 & 1+j5 \\ 2+j & j & 3-j \\ 3 & 1 & 1+j3 \end{bmatrix}$$

então

$$\bar{\mathbf{A}}' = \mathbf{A}^* = \begin{bmatrix} 1 & 2-j & 3 \\ -j2 & -j & 1 \\ 1-j5 & 3+j & 1-j3 \end{bmatrix}$$

Note que

$$(\mathbf{A}^*)^* = \mathbf{A}$$

Se **A** é uma matriz real (uma matriz cujos elementos são reais), a transposta conjugada $\mathbf{A}^*$ é igual à transposta $\mathbf{A}'$.

Matriz Hermitiana. Uma matriz cujos elementos são grandezas complexas é denominada uma matriz *complexa*. Se uma matriz complexa **A** satisfaz a relação

$$\mathbf{A} = \mathbf{A}^* \quad \text{ou} \quad a_{ij} = \bar{a}_{ji}$$

onde $\bar{a}_{ji}$ é o complexo conjugado de a_{ji}, então **A** é denominada matriz *Hermitiana*. Um exemplo é

$$\mathbf{A} = \begin{bmatrix} 1 & 4+j3 \\ 4-j3 & 2 \end{bmatrix}$$

Se uma matriz Hermitiana **A** é escrita como $\mathbf{A} = \mathbf{B} + j\mathbf{C}$, onde **B** e **C** são matrizes reais, então

$$\mathbf{B} = \mathbf{B}' \quad \text{e} \quad \mathbf{C} = -\mathbf{C}'$$

No exemplo acima,

$$\mathbf{A} = \mathbf{B} + j\mathbf{C} = \begin{bmatrix} 1 & 4 \\ 4 & 2 \end{bmatrix} + j\begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}$$

Matriz anti-Hermitiana. Se uma matriz **A** satisfaz a relação

$$\mathbf{A} = -\mathbf{A}^*$$

então **A** é denominada matriz *anti-Hermitiana*. Um exemplo é

$$\mathbf{A} = \begin{bmatrix} j5 & -2+j3 \\ 2+j3 & j \end{bmatrix}$$

Se uma matriz anti-Hermitiana **A** é escrita como **A** = **B** + j**C**, onde **B** e **C** são matrizes reais, então

$$\mathbf{B} = -\mathbf{B}' \qquad \text{e} \qquad \mathbf{C} = \mathbf{C}'$$

No presente exemplo,

$$\mathbf{A} = \mathbf{B} + j\mathbf{C} = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix} + j\begin{bmatrix} 5 & 3 \\ 3 & 1 \end{bmatrix}$$

3.3 ÁLGEBRA MATRICIAL

Esta seção apresenta a álgebra matricial, bem como definições adicionais. É importante lembrar que algumas das operações matriciais obedecem às mesmas regras da álgebra comum, porém o mesmo não ocorre com outras operações matriciais.

Adição e subtração de matrizes. Duas matrizes **A** e **B** podem ser adicionadas se possuírem o mesmo número de linhas e o mesmo número de colunas. Se **A** = (a_{ij}) e **B** = (b_{ij}), então **A** + **B** é definida como

$$\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$$

Portanto, cada elemento de **A** é adicionado ao elemento correspondente de **B**. De modo equivalente, a subtração de matrizes é definida como

$$\mathbf{A} - \mathbf{B} = (a_{ij} - b_{ij})$$

Como um exemplo, considere

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad \text{e} \qquad \mathbf{B} = \begin{bmatrix} 5 & 2 & 3 \\ 1 & 4 & 1 \end{bmatrix}$$

Então **A** + **B** e **A** − **B** são dadas por

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 6 & 4 & 6 \\ 5 & 9 & 7 \end{bmatrix} \qquad \text{e} \qquad \mathbf{A} - \mathbf{B} = \begin{bmatrix} -4 & 0 & 0 \\ 3 & 1 & 5 \end{bmatrix}$$

Multiplicação de uma matriz por um escalar. O produto de uma matriz por um escalar é uma matriz na qual cada elemento é multiplicado pelo escalar; isto é, para uma matriz **A** e um escalar k,

$$k\mathbf{A} = \begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1m} \\ ka_{21} & ka_{22} & \cdots & ka_{2m} \\ \vdots & \vdots & & \vdots \\ ka_{n1} & ka_{n2} & \cdots & ka_{nm} \end{bmatrix}$$

Multiplicação de uma matriz por uma matriz. Multiplicação de uma matriz por uma matriz é possível entre matrizes conformes (o que significa que o número de colunas da primeira matriz deve ser igual ao número de linhas da segunda). Caso contrário, não é definida a multiplicação de duas matrizes.

Seja **A** uma matriz $n \times m$ e **B** uma matriz $m \times p$. Então, o produto **AB**, que deve ser lido "**A** pós-multiplicada por **B**" ou "**B** pré-multiplicada por **A**", é definido como segue:

$$\mathbf{AB} = \mathbf{C} = (c_{ij}) = \left(\sum_{k=1}^{m} a_{ik}b_{kj}\right) \qquad (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p)$$

A matriz produto **C** possui o mesmo número de linhas da matriz **A** e o mesmo número de colunas da matriz **B**. Conseqüentemente, a matriz **C** é uma matriz $n \times p$.

Deve-se notar que mesmo que **A** e **B** sejam conformes para o produto **AB**, elas podem não ser conformes para **BA**, quando, então, **BA** não é definida.

As leis associativa e distributiva são aplicáveis para a multiplicação matricial; isto é,

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$
$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$$
$$\mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB}$$

Se **AB** = **BA**, então **A** e **B** são ditas comutarem. Note que, em geral, **AB** ≠ **BA**. Para mostrar este fato, considere

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 4 \end{bmatrix} \qquad \text{e} \qquad \mathbf{B} = \begin{bmatrix} 2 & 1 & 5 \\ 1 & 0 & 0 \end{bmatrix}$$

AB e **BA** são dadas, respectivamente, por

$$\mathbf{AB} = \begin{bmatrix} 2 & 1 & 5 \\ 5 & 2 & 10 \\ 10 & 3 & 15 \end{bmatrix} \qquad \text{e} \qquad \mathbf{BA} = \begin{bmatrix} 19 & 21 \\ 1 & 0 \end{bmatrix}$$

Obviamente, **AB** ≠ **BA**. Para um outro exemplo, considere

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \qquad \text{e} \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$$

Então,

$$\mathbf{AB} = \begin{bmatrix} 5 & 6 \\ 10 & 12 \end{bmatrix} \quad \text{e} \quad \mathbf{BA} = \begin{bmatrix} 5 & 10 \\ 6 & 12 \end{bmatrix}$$

Obviamente, $\mathbf{AB} \neq \mathbf{BA}$.

Como a multiplicação matricial, em geral, é não comutativa, devemos preservar a ordem das matrizes quando multiplicarmos uma matriz por uma outra matriz. (Esta é a razão por que sempre usamos os termos "pré-multiplicação" ou "pós-multiplicação" para indicar quando a matriz é multiplicada pela direita ou pela esquerda.)

Um exemplo do caso onde $\mathbf{AB} = \mathbf{BA}$ é dado abaixo:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$$

AB e **BA** são dadas por

$$\mathbf{AB} = \mathbf{BA} = \begin{bmatrix} 2 & 0 \\ 0 & 15 \end{bmatrix}$$

É claro que **A** e **B** comutam neste caso.

Potência de uma matriz. A k-ésima potência de uma matriz quadrada **A** é definida por

$$\mathbf{A}^k = \underbrace{\mathbf{AA} \cdots \mathbf{A}}_{k}$$

Note que para uma matriz diagonal $\mathbf{A} = \text{diag}(a_{11}, a_{22}, \ldots, a_{nn})$.

$$\mathbf{A}^k = \begin{bmatrix} a_{11}^k & & & 0 \\ & a_{22}^k & & \\ & & \ddots & \\ 0 & & & a_{nn}^k \end{bmatrix} = \text{diag}\,(a_{11}^k, a_{22}^k, \ldots, a_{nn}^k)$$

Outras propriedades de matrizes. As transpostas de $\mathbf{A} + \mathbf{B}$ e **AB** são dadas por

$$(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$$
$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

Para provar a última relação, note que o (i, j)ésimo elemento de **AB** é

$$\sum_{k=1}^{m} a_{ik} b_{kj} = c_{ij}$$

O (i, j)ésimo elemento de $\mathbf{B}'\mathbf{A}'$ é

$$\sum_{k=1}^{m} b_{ki} a_{jk} = \sum_{k=1}^{m} a_{jk} b_{ki} = c_{ji}$$

que é igual ao (j, i)ésimo elemento de **AB**, ou o (i, j)ésimo elemento de $(\mathbf{AB})'$. Portanto, $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$. Como um exemplo, considere

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \quad \text{e} \quad \mathbf{B} = \begin{bmatrix} 4 & 2 \\ 5 & 6 \end{bmatrix}$$

então,

$$\mathbf{AB} = \begin{bmatrix} 24 & 26 \\ 23 & 22 \end{bmatrix}$$

$$\mathbf{B}'\mathbf{A}' = \begin{bmatrix} 4 & 5 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 24 & 23 \\ 26 & 22 \end{bmatrix}$$

Obviamente, $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.

De um modo similar, obtemos as transpostas conjugadas de $\mathbf{A} + \mathbf{B}$ e **AB**, resultando

$$(\mathbf{A} + \mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$$
$$(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$$

A característica (*rank*) de uma matriz. Uma matriz é dita possuir uma característica m se existir uma submatriz $\mathbf{M}$ $m \times m$ de **A** de modo que o determinante de **M** é não nulo e o determinante de todas as submatrizes $r \times r$ (onde $r \geq m + 1$) de **A** é zero.

Como um exemplo, considere a seguinte matriz:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 2 \end{bmatrix}$$

Note que $|\mathbf{A}| = 0$. Uma das submatrizes maiores cujo determinante não é igual a zero é a seguinte:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

Portanto a característica da matriz **A** é 3.

3.4 INVERSÃO DE MATRIZ

Esta seção discute a inversão de matrizes e tópicos associados.

Menor M_{ij}. Se a i-ésima linha e a j-ésima coluna são eliminadas de uma matriz **A** $n \times n$, a matriz resultante é uma matriz $(n-1) \times (n-1)$. O determinante desta matriz $(n-1) \times (n-1)$ é denominado o menor M_{ij} da matriz **A**.

Cofator A_{ij}. O cofator A_{ij} do elemento a_{ij} da matriz **A** é definido pela equação

$$A_{ij} = (-1)^{i+j} M_{ij}$$

Isto é, o cofator A_{ij} do elemento a_{ij} é $(-1)^{i+j}$ vezes o determinante da matriz constituída pela eliminação da i-ésima linha e da j-ésima coluna de **A**. Note que o cofator A_{ij} do elemento a_{ij} é o coeficiente do termo a_{ij} na expansão do determinante $|\mathbf{A}|$, desde que pode ser demonstrado que

$$a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = |\mathbf{A}|$$

Se $a_{i1}, a_{i2}, \ldots, a_{in}$ são substituídos por $a_{j1}, a_{j2}, \ldots, a_{jn}$, então

$$a_{j1}A_{i1} + a_{j2}A_{i2} + \cdots + a_{jn}A_{in} = 0 \qquad (i \neq j)$$

desde que o determinante de **A** neste caso possui duas linhas idênticas. Conseqüentemente, obtemos

$$\sum_{k=1}^{n} a_{jk}A_{ik} = \delta_{ji}|\mathbf{A}|$$

Analogamente,

$$\sum_{k=1}^{n} a_{ki}A_{kj} = \delta_{ij}|\mathbf{A}|$$

Matriz adjunta. A matriz **B**, onde o elemento da i-ésima linha e j-ésima coluna é igual a A_{ji}, é denominada a adjunta de **A** e é denotada por adj **A** ou

$$\mathbf{B} = (b_{ij}) = (A_{ji}) = \text{adj } \mathbf{A}$$

Isto é, a adjunta de **A** é a transposta da matriz cujos elementos são os cofatores de **A** ou

$$\text{adj } \mathbf{A} = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

Note que o elemento da j-ésima linha e i-ésima coluna do produto **A**(adj **A**) é

$$\sum_{k=1}^{n} a_{jk}b_{ki} = \sum_{k=1}^{n} a_{jk}A_{ik} = \delta_{ji}|\mathbf{A}|$$

Então, **A**(adj **A**) é uma matriz diagonal cujos elementos da diagonal são iguais a $|\mathbf{A}|$.

Portanto,

$$\mathbf{A}(\text{adj } \mathbf{A}) = |\mathbf{A}|\mathbf{I}$$

Analogamente, o elemento na j-ésima linha e i-ésima coluna do produto (adj **A**)**A** é

$$\sum_{k=1}^{n} b_{jk}a_{ki} = \sum_{k=1}^{n} A_{kj}a_{ki} = \delta_{ij}|\mathbf{A}|$$

Conseqüentemente, temos a relação

$$\mathbf{A}(\text{adj } \mathbf{A}) = (\text{adj } \mathbf{A})\mathbf{A} = |\mathbf{A}|\mathbf{I} \tag{3-1}$$

Por exemplo, dada a matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & -2 \\ 1 & 0 & -3 \end{bmatrix}$$

podemos verificar que o determinante de **A** é 17 e que

$$\text{adj } \mathbf{A} = \begin{bmatrix} \begin{vmatrix} -1 & -2 \\ 0 & -3 \end{vmatrix} & -\begin{vmatrix} 2 & 0 \\ 0 & -3 \end{vmatrix} & \begin{vmatrix} 2 & 0 \\ -1 & -2 \end{vmatrix} \\ -\begin{vmatrix} 3 & -2 \\ 1 & -3 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 1 & -3 \end{vmatrix} & -\begin{vmatrix} 1 & 0 \\ 3 & -2 \end{vmatrix} \\ \begin{vmatrix} 3 & -1 \\ 1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 6 & -4 \\ 7 & -3 & 2 \\ 1 & 2 & -7 \end{bmatrix}$$

Portanto

$$\mathbf{A}(\text{adj } \mathbf{A}) = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & -2 \\ 1 & 0 & -3 \end{bmatrix} \begin{bmatrix} 3 & 6 & -4 \\ 7 & -3 & 2 \\ 1 & 2 & -7 \end{bmatrix}$$

$$= 17 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= |\mathbf{A}|\mathbf{I}$$

Inversa de uma matriz. Se, para uma matriz quadrada **A**, existir uma matriz **B** tal que **BA** = **AB** = **I**, então **B** é denotada por $\mathbf{A}^{-1}$ e é denominada a *inversa* de **A**. A

inversa de uma matriz A existe se o determinante de A é não nulo ou A é não singular.

Por definição, a matriz inversa $\mathbf{A}^{-1}$ possui a propriedade de que

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

onde I é a matriz identidade. Se A é não singular e AB = C, então $\mathbf{B} = \mathbf{A}^{-1}\mathbf{C}$. Isto pode ser visto da equação

$$\mathbf{A}^{-1}\mathbf{A}\mathbf{B} = \mathbf{I}\mathbf{B} = \mathbf{B} = \mathbf{A}^{-1}\mathbf{C}$$

Se A e B são matrizes não singulares, então o produto AB é uma matriz não singular. Além disso,

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

Isto pode ser provado como segue:

$$(\mathbf{B}^{-1}\mathbf{A}^{-1})\mathbf{A}\mathbf{B} = \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}\mathbf{I}\mathbf{B} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$$

Analogamente,

$$(\mathbf{A}\mathbf{B})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{I}$$

Note que

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$$

$$(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}$$

$$(\mathbf{A}^{-1})^* = (\mathbf{A}^*)^{-1}$$

Da Eq. (3.1) e da definição de matriz inversa, temos

$$\mathbf{A}^{-1} = \frac{\text{adj } \mathbf{A}}{|\mathbf{A}|}$$

Conseqüentemente, a inversa de uma matriz é a transposta da matriz de seus cofatores, dividida pelo determinante da matriz. Se A é dada por

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

então

$$\mathbf{A}^{-1} = \frac{\text{adj } \mathbf{A}}{|\mathbf{A}|} = \begin{bmatrix} \frac{A_{11}}{|\mathbf{A}|} & \frac{A_{21}}{|\mathbf{A}|} & \cdots & \frac{A_{n1}}{|\mathbf{A}|} \\ \frac{A_{12}}{|\mathbf{A}|} & \frac{A_{22}}{|\mathbf{A}|} & \cdots & \frac{A_{n2}}{|\mathbf{A}|} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{A_{1n}}{|\mathbf{A}|} & \frac{A_{2n}}{|\mathbf{A}|} & \cdots & \frac{A_{nn}}{|\mathbf{A}|} \end{bmatrix}$$

onde A_{ij} é o cofator de a_{ij} da matriz A. Portanto, os termos na i-ésima coluna de $\mathbf{A}^{-1}$ são $1/|\mathbf{A}|$ vezes os cofatores da i-ésima linha da matriz original A. Por exemplo, se A é dada por

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & -2 \\ 1 & 0 & -3 \end{bmatrix}$$

então a adjunta de A e o determinante $|\mathbf{A}|$ são os seguintes:

$$\text{adj } \mathbf{A} = \begin{bmatrix} 3 & 6 & -4 \\ 7 & -3 & 2 \\ 1 & 2 & -7 \end{bmatrix}, \qquad |\mathbf{A}| = 17$$

Então a inversa de A é dada por

$$\mathbf{A}^{-1} = \frac{\text{adj } \mathbf{A}}{|\mathbf{A}|} = \begin{bmatrix} \frac{3}{17} & \frac{6}{17} & -\frac{4}{17} \\ \frac{7}{17} & -\frac{3}{17} & \frac{2}{17} \\ \frac{1}{17} & \frac{2}{17} & -\frac{7}{17} \end{bmatrix}$$

A seguir fornecemos fórmulas para determinar matrizes inversas para matrizes 2×2 e matrizes 3×3. Para a matriz A 2×2, onde

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad ad - bc \neq 0$$

a matriz inversa é dada por

$$\mathbf{A}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Para a matriz A 3×3, onde

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}, \qquad |\mathbf{A}| \neq 0$$

a matriz inversa é dada por

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\begin{bmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} b & c \\ e & f \end{vmatrix} \\ -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} \\ \begin{vmatrix} d & e \\ g & h \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{bmatrix}$$

Observações sobre cancelamentos de matrizes. Cancelamento de matrizes não é válido em álgebra matricial. Considere o produto de duas matrizes singulares **A** e **B**. Por exemplo,

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix} \neq \mathbf{0} \qquad \text{e} \qquad \mathbf{B} = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \neq \mathbf{0}$$

Então

$$\mathbf{AB} = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0}$$

Obviamente, $\mathbf{AB} = \mathbf{0}$ não implica nem $\mathbf{A} = \mathbf{0}$ nem $\mathbf{B} = \mathbf{0}$. De fato, $\mathbf{AB} = \mathbf{0}$ implica uma das três hipóteses seguintes:

1. $\mathbf{A} = \mathbf{0}$
2. $\mathbf{B} = \mathbf{0}$
3. Tanto **A** como **B** são singulares.

Podemos provar facilmente que, se tanto **A** como **B** são matrizes não nulas e $\mathbf{AB} = \mathbf{0}$, então tanto **A** como **B** são singulares. Para uma prova, suponha que se **A** e **B** não são singulares, então existe uma matriz $\mathbf{A}^{-1}$ com a seguinte propriedade:

$$\mathbf{A}^{-1}\mathbf{AB} = \mathbf{B} = \mathbf{0}$$

o que contradiz a hipótese de que **B** é uma matriz não nula. Portanto, concluímos que tanto **A** como **B** devem ser singulares se $\mathbf{A} \neq \mathbf{0}$ e $\mathbf{B} \neq \mathbf{0}$.

Analogamente, note que se **A** é singular, então nem $\mathbf{AB} = \mathbf{AC}$ nem $\mathbf{BA} = \mathbf{CA}$ implica $\mathbf{B} = \mathbf{C}$. Se, entretanto, **A** é uma matriz não singular, então $\mathbf{AB} = \mathbf{AC}$ implica $\mathbf{B} = \mathbf{C}$ e $\mathbf{BA} = \mathbf{CA}$ também implica $\mathbf{B} = \mathbf{C}$.

3.5 DIFERENCIAÇÃO E INTEGRAÇÃO DE MATRIZES

A derivada de uma matriz $\mathbf{A}(t)$ $n \times m$ é definida como a matriz $n \times m$ cujos elementos correspondem à derivada dos elementos correspondentes da matriz original, considerando-se que todos os elementos $a_{ij}(t)$ possuem derivadas em relação a t. Isto é,

$$\frac{d}{dt}\mathbf{A}(t) = \left(\frac{d}{dt}a_{ij}(t)\right) = \begin{bmatrix} \frac{d}{dt}a_{11}(t) & \frac{d}{dt}a_{12}(t) & \cdots & \frac{d}{dt}a_{1m}(t) \\ \frac{d}{dt}a_{21}(t) & \frac{d}{dt}a_{22}(t) & \cdots & \frac{d}{dt}a_{2m}(t) \\ \vdots & \vdots & & \vdots \\ \frac{d}{dt}a_{n1}(t) & \frac{d}{dt}a_{n2}(t) & \cdots & \frac{d}{dt}a_{nm}(t) \end{bmatrix}$$

Analogamente, a integral de uma matriz $\mathbf{A}(t)$ $n \times m$ é definida por

$$\int \mathbf{A}(t)\,dt = \left(\int a_{ij}(t)\,dt\right) = \begin{bmatrix} \int a_{11}(t)\,dt & \int a_{12}(t)\,dt & \cdots & \int a_{1m}(t)\,dt \\ \int a_{21}(t)\,dt & \int a_{22}(t)\,dt & \cdots & \int a_{2m}(t)\,dt \\ \vdots & \vdots & & \vdots \\ \int a_{n1}(t)\,dt & \int a_{n2}(t)\,dt & \cdots & \int a_{nm}(t)\,dt \end{bmatrix}$$

Diferenciação do produto de duas matrizes. Se as matrizes $\mathbf{A}(t)$ e $\mathbf{B}(t)$ podem ser diferenciadas em relação a t, então

$$\frac{d}{dt}[\mathbf{A}(t)\mathbf{B}(t)] = \frac{d\mathbf{A}(t)}{dt}\mathbf{B}(t) + \mathbf{A}(t)\frac{d\mathbf{B}(t)}{dt}$$

Novamente, neste caso a multiplicação de $\mathbf{A}(t)$ e $d\mathbf{B}(t)/dt$ [ou $d\mathbf{A}(t)/dt$ e $\mathbf{B}(t)$] é, em geral, não comutativa.

Diferenciação de $\mathbf{A}^{-1}(t)$. Se uma matriz $\mathbf{A}(t)$ e sua inversa $\mathbf{A}^{-1}(t)$ são diferenciáveis em relação a t, então a derivada de $\mathbf{A}^{-1}(t)$ é dada por

$$\frac{d\mathbf{A}^{-1}(t)}{dt} = -\mathbf{A}^{-1}(t)\frac{d\mathbf{A}(t)}{dt}\mathbf{A}^{-1}(t)$$

Este resultado pode ser demonstrado pela diferenciação de $\mathbf{A}(t)\mathbf{A}^{-1}(t)$ em relação a t. Desde que

$$\frac{d}{dt}[\mathbf{A}(t)\mathbf{A}^{-1}(t)] = \frac{d\mathbf{A}(t)}{dt}\mathbf{A}^{-1}(t) + \mathbf{A}(t)\frac{d\mathbf{A}^{-1}(t)}{dt}$$

e

$$\frac{d}{dt}\mathbf{A}(t)\mathbf{A}^{-1}(t) = \frac{d}{dt}\mathbf{I} = \mathbf{0}$$

obtemos

$$\mathbf{A}(t)\frac{d\mathbf{A}^{-1}(t)}{dt} = -\frac{d\mathbf{A}(t)}{dt}\mathbf{A}^{-1}(t)$$

ou

$$\frac{d\mathbf{A}^{-1}(t)}{dt} = -\mathbf{A}^{-1}(t)\frac{d\mathbf{A}(t)}{dt}\mathbf{A}^{-1}(t)$$

PROBLEMAS ILUSTRATIVOS E SOLUÇÕES

Problema A.3.1 Mostre que se **A** é uma matriz quadrada qualquer, então **A** + **A**′ é uma matriz simétrica e **A** − **A**′ é uma matriz anti-simétrica.

Solução. Considere a seguinte matriz 3 × 3:

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Então

$$\mathbf{A} + \mathbf{A}' = \begin{bmatrix} 2a & b+d & c+g \\ b+d & 2e & f+h \\ c+g & f+h & 2i \end{bmatrix} = \text{matriz simétrica}$$

e

$$\mathbf{A} - \mathbf{A}' = \begin{bmatrix} 0 & b-d & c-g \\ d-b & 0 & f-h \\ g-c & h-f & 0 \end{bmatrix} = \text{matriz anti-simétrica}$$

Problema A.3.2 Calcule o produto **AB**, onde

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 1 \\ 1 & 5 & 2 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 1 & -1 \end{bmatrix}$$

Solução.

$$\mathbf{AB} = \begin{bmatrix} 5 & 6 \\ 2 & -4 \\ 13 & 13 \end{bmatrix}$$

Problema A.3.3 A matriz de Vandermonde é dada por

$$\mathbf{V} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{bmatrix}$$

Determine o determinante de **V** para os casos em que $n = 3$ e $n = 4$.

Solução. Para $n = 3$,

$$|\mathbf{V}| = (x_3 - x_2)(x_3 - x_1)(x_2 - x_1)$$

Para $n = 4$,

$$|\mathbf{V}| = (x_4 - x_3)(x_4 - x_2)(x_4 - x_1)(x_3 - x_2)(x_3 - x_1)(x_2 - x_1)$$

Problema A.3.4 Demonstre que se as matrizes **A** e **B** são $n \times n$, então,

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{vmatrix} = |\mathbf{A} + \mathbf{B}| \cdot |\mathbf{A} - \mathbf{B}|$$

Solução.

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{vmatrix} = \begin{vmatrix} \mathbf{A} + \mathbf{B} & \mathbf{B} + \mathbf{A} \\ \mathbf{B} & \mathbf{A} \end{vmatrix} = \begin{vmatrix} \mathbf{A} + \mathbf{B} & \mathbf{0} \\ \mathbf{B} & \mathbf{A} - \mathbf{B} \end{vmatrix} = |\mathbf{A} + \mathbf{B}| \cdot |\mathbf{A} - \mathbf{B}|$$

Note que para as matrizes **A**, **B**, **C** e **D**, $n \times n$, o determinante

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix}$$

é, em geral, diferente de $|\mathbf{AD} - \mathbf{BC}|$ e de $|\mathbf{A}|\,|\mathbf{D}| - |\mathbf{B}|\,|\mathbf{C}|$.

Problema A.3.5 Determine a inversa de

$$\mathbf{A} = \begin{bmatrix} s+1 & -1 \\ s+2 & s \end{bmatrix}$$

onde s é uma variável complexa.

Solução.

$$\mathbf{A}^{-1} = \begin{bmatrix} \dfrac{s}{s^2 + 2s + 2} & \dfrac{1}{s^2 + 2s + 2} \\ -\dfrac{s+2}{s^2 + 2s + 2} & \dfrac{s-1}{s^2 + 2s + 2} \end{bmatrix}$$

Problema A.3.6 Determine a solução do seguinte conjunto de três equações simultâneas:

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 1 \\ 2x_1 + x_2 &= 2 \\ x_1 + x_2 + 2x_3 &= 3 \end{aligned}$$

Solução. Definindo as seguintes matrizes e vetores:

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

obtemos

$$\mathbf{Ax} = \mathbf{b}$$

Portanto,

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

onde

$$\mathbf{A}^{-1} = \frac{1}{-9}\begin{bmatrix} \begin{vmatrix}1 & 0\\1 & 2\end{vmatrix} & -\begin{vmatrix}3 & 1\\1 & 2\end{vmatrix} & \begin{vmatrix}3 & 1\\1 & 0\end{vmatrix} \\ -\begin{vmatrix}2 & 0\\1 & 2\end{vmatrix} & \begin{vmatrix}1 & 1\\1 & 2\end{vmatrix} & -\begin{vmatrix}1 & 1\\2 & 0\end{vmatrix} \\ \begin{vmatrix}2 & 1\\1 & 1\end{vmatrix} & -\begin{vmatrix}1 & 3\\1 & 1\end{vmatrix} & \begin{vmatrix}1 & 3\\2 & 1\end{vmatrix} \end{bmatrix} = \begin{bmatrix} -\frac{2}{9} & \frac{5}{9} & \frac{1}{9} \\ \frac{4}{9} & -\frac{1}{9} & -\frac{2}{9} \\ -\frac{1}{9} & -\frac{2}{9} & \frac{5}{9} \end{bmatrix}$$

Resultando

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{2}{9} & \frac{5}{9} & \frac{1}{9} \\ \frac{4}{9} & -\frac{1}{9} & -\frac{2}{9} \\ -\frac{1}{9} & -\frac{2}{9} & \frac{5}{9} \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{11}{9} \\ -\frac{4}{9} \\ \frac{10}{9} \end{bmatrix}$$

PROBLEMAS

Problema B.3.1 Demonstre que qualquer matriz quadrada pode ser escrita como sendo a soma de uma matriz Hermitiana e outra anti-Hermitiana.

Problema B.3.2 Demonstre que para as matrizes **A** e **B**:

1. $|\mathbf{A}| = |\mathbf{A}'|$
2. $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$
3. $|(\mathbf{AB})'| = |\mathbf{A}||\mathbf{B}|$

Problema B.3.3 Em geral, $(\mathbf{A} + \mathbf{B})^2$ não é igual a $\mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$. Explique a razão. Sob qual condição $(\mathbf{A} + \mathbf{B})^2$ é igual a $\mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$?

Problema B.3.4 Determine $\mathbf{A}^{-1}$ considerando que

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}$$

Problema B.3.5 Uma matriz cujos elementos abaixo (ou acima) da diagonal principal são todos nulos é denominada uma matriz triangular. Um exemplo é

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$$

Usando esta matriz 3 × 3, demonstre que a matriz inversa de uma matriz triangular também é uma matriz triangular.

Problema B.3.6 Determine a característica da seguinte matriz:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 & 0 \\ 2 & 0 & 1 & 3 \\ 4 & -1 & 5 & 1 \end{bmatrix}$$

Problema B.3.7 Resolva o seguinte conjunto de três equações simultâneas:

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 2 \\ x_1 - 2x_2 + x_3 &= 3 \\ 2x_1 + x_2 + 3x_3 &= 1 \end{aligned}$$

Problema B.3.8 Um conjunto de n equações algébricas simultâneas:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots \quad & \quad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

pode ser descrito por

$$\mathbf{Ax} = \mathbf{b}$$

onde

$$\mathbf{A} = (a_{ij})$$

Um sistema de equações é dito *consistente* se possuir pelo menos uma solução. Se as n equações são consistentes porém **A** é singular, como podemos determinar **x**? Utilizando um caso simples onde $n = 3$, explique como se pode obter **x**. Considere os dois casos nos quais a característica de $\mathbf{A} = 2$ e a característica de $\mathbf{A} = 1$.

4

Modelos Matemáticos de Sistemas Físicos

4.1 INTRODUÇÃO

A maioria dos sistemas dinâmicos, independentemente de serem mecânicos, elétricos, térmicos, hidráulicos, econômicos, biológicos etc., podem ser caracterizados por equações diferenciais. A resposta de um sistema dinâmico a uma dada entrada (ou função de excitação) pode ser obtida se estas equações diferenciais são resolvidas. Podem-se obter as equações utilizando leis físicas que governam um particular sistema, por exemplo, as leis de Newton para sistemas mecânicos, as leis de Kirchhoff para sistemas elétricos etc.

Modelos matemáticos. A descrição matemática das características dinâmicas de um sistema é denominado *modelo matemático*. O primeiro passo na análise de um sistema dinâmico é obter seu modelo. Deve-se sempre levar em conta que a obtenção de um modelo matemático razoável é a parte mais importante de toda a análise.

Os modelos podem assumir formas muito diferentes. Dependendo do particular sistema e de certas circunstâncias, uma representação matemática pode ser mais conveniente do que outras representações. Por exemplo, em problemas de controle ótimo, é quase sempre vantajoso usar um conjunto de equações diferenciais de primeira-ordem. (Referir-se ao Cap. 14.) Por outro lado, para a análise de resposta transitória ou análise de resposta em freqüência de sistemas de entrada-simples-saída-simples, a representação através da função de transferência a ser discutida neste capítulo poderá ser mais conveniente do que qualquer outra.

Uma vez obtido o modelo matemático de um sistema, várias ferramentas analíticas ou por computador podem ser utilizadas para fins de análise e síntese.

Simplicidade *versus* precisão. Na obtenção de um modelo, devemos estabelecer um compromisso entre a simplicidade do modelo e a precisão dos resultados da análise. Note que os resultados obtidos da análise são válidos somente na extensão da qual o modelo se aproxima de um dado sistema físico.

A rapidez com a qual um computador digital pode desempenhar operações aritméticas permite-nos empregar uma nova abordagem na formulação de modelos matemáticos. Ao invés de nos limitarmos apenas aos modelos simples, podemos, se necessário, incluir centenas de equações para descrever um sistema completo. Se não for necessário uma precisão extrema, entretanto, é preferível obter apenas um modelo razoavelmente simplificado.

Na dedução de tal modelo simplificado, freqüentemente torna-se necessário ignorar certas propriedades físicas inerentes ao sistema. Em particular, se um modelo matemático linear a parâmetros concentrados (isto é, um modelo empregando equações diferenciais ordinárias) é desejado, é sempre necessário ignorar certas não linearidades e parâmetros distribuídos (isto é, aqueles que resultam em equações diferenciais parciais) que podem estar presentes no sistema físico. Se os efeitos que estas propriedades ignoradas apresentam na resposta são pequenos, será obtida uma boa concordância entre os resultados da análise de um modelo matemático e os resultados do estudo experimental do sistema físico.

Em geral, na solução de um novo problema, verificamos ser desejável inicialmente construir um modelo simplificado de modo a ganharmos um conhecimento básico e geral para a solução. Posteriormente um modelo matemático mais completo poderá então ser elaborado e utilizado para uma análise mais integral.

Devemos estar cientes do fato de que um modelo linear a parâmetros concentrados, que pode ser válido em operações de baixa freqüência, poderá não ser válido em freqüências suficientemente altas desde que as propriedades desprezadas dos parâmetros distribuídos podem tornar-se um fator importante no comportamento dinâmico do sistema. Por exemplo, a massa de uma mola pode ser desprezada em operações de baixa freqüência, porém torna-se uma propriedade importante do sistema em altas freqüências.

Sistemas lineares. Sistemas lineares são aqueles nos quais as equações do modelo são lineares. Uma equação diferencial é linear se os coeficientes são constantes ou apenas funções da variável independente. A propriedade mais importante de sistemas lineares é que o princípio da superposição é aplicável. O princípio da superposição estabelece que a resposta produzida pela aplicação simultânea de duas forças de excitação diferentes é igual à soma das duas respostas individuais. Conseqüentemente, para sistemas lineares, a resposta para várias entradas pode ser calculada considerando-se uma única entrada de cada vez e adicionando-se os resultados. É este princípio que permite a construção de soluções complicadas para equações diferenciais lineares a partir de soluções relativamente simples.

Em uma investigação experimental de um sistema dinâmico, se a causa e o efeito são proporcionais, implicando que o princípio da superposição é aplicável, então o sistema pode ser considerado linear.

Sistemas lineares invariáveis no tempo e sistemas lineares variáveis no tempo. Sistemas dinâmicos que são lineares e compostos de componentes com parâmetros concentrados e invariáveis no tempo podem ser descritos por equações diferenciais invariáveis no tempo. Estes sistemas são denominados sistemas *lineares invariáveis no tempo* (a *coeficiente-constante linear*). Os sistemas que são representados por equações diferenciais cujos coeficientes são funções de tempo são denominados sistemas variáveis no tempo. Um exemplo de um sistema de controle variável no tempo é um sistema de controle de uma nave espacial. (A

massa de uma nave espacial varia devido ao consumo de combustível, e a força gravitacional varia conforme a nave se distancia da Terra.)

Sistemas não lineares. Sistemas não lineares são aqueles representados por equações não lineares. Exemplos de equações não lineares são

$$y = \operatorname{sen} x$$

$$y = x^2$$

$$z = x^2 + y^3$$

(Na última equação, z é uma função não linear de x e y.)

Uma equação diferencial é dita *não linear* se ela não for linear. Exemplos de equações diferenciais não lineares são

$$\frac{d^2x}{dt^2} + \left(\frac{dx}{dt}\right)^2 + x = A \operatorname{sen} \omega t$$

$$\frac{d^2x}{dt^2} + (x^2 - 1)\frac{dx}{dt} + x = 0$$

$$\frac{d^2x}{dt^2} + \frac{dx}{dt} + x + x^3 = 0$$

Embora muitas relações físicas sejam representadas muitas vezes por equações lineares, na maioria dos casos as relações reais não são exatamente lineares. De fato, um estudo cuidadoso de sistemas físicos revela que mesmo os chamados "sistemas lineares" são realmente lineares apenas em faixas limitadas de operação. Na prática, muitos sistemas eletromecânicos, sistemas hidráulicos, sistemas pneumáticos etc. envolvem relações não lineares entre as variáveis. Por exemplo, a saída de um componente pode saturar para sinais grandes na entrada. Pode haver, por outro lado, um espaço morto que afeta em relação a pequenos sinais. (O espaço morto de um componente é um pequeno intervalo de variações na entrada dentro do qual o componente é insensível.) Não linearidade do tipo lei-quadrática pode ocorrer em alguns componentes. Por exemplo, amortecedores utilizados em sistemas físicos podem ser lineares em operações de baixa velocidade, porém podem tornar-se não lineares para altas velocidades, e a força amortecedora pode tornar-se proporcional ao quadrado da velocidade de operação. Exemplos de curvas características para estas não linearidades são indicados na Fig. 4.1.

Note que alguns sistemas de controle importantes são não lineares em relação a sinais de qualquer dimensão. Por exemplo, em sistemas de controle liga-desliga (*on-off*), a ação de controle ou é liga ou desliga e não há uma relação linear entre a entrada e a saída do controlador.

A característica mais importante de sistemas não lineares é o fato de que não é aplicável o princípio da superposição. Os procedimentos para determinar as soluções de problemas que possuam sistemas não lineares, em geral, são extremamente complicados. Devido a esta dificuldade matemática inerente a sistemas não lineares, normalmente é necessário introduzir sistemas lineares "equivalentes" no lugar daqueles não lineares. Estes sistemas lineares equivalentes somente são válidos dentro de uma faixa limitada de operação. Uma vez que um sistema não linear é aproximado por um modelo matemático linear, várias ferramentas lineares podem

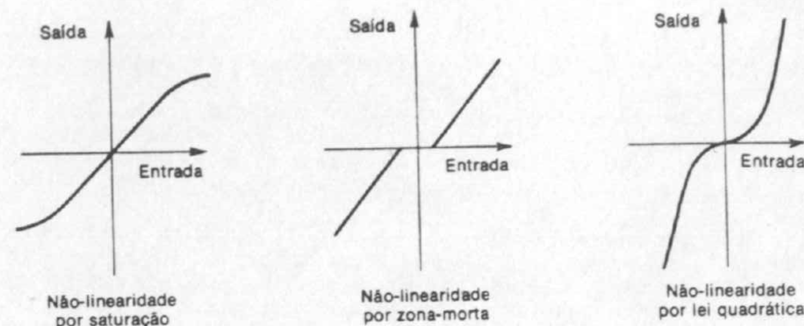


Fig. 4.1 Curvas características para várias não-linearidades.

ser aplicadas para fins de análise e projeto. Introduziremos neste livro várias técnicas de linearização.

4.2 FUNÇÕES DE TRANSFERÊNCIA

Na teoria de controle, funções denominadas "funções de transferência" são extremamente usadas para caracterizar as relações entrada-saída de sistemas lineares invariáveis no tempo. O conceito de funções de transferência aplica-se somente a sistemas lineares invariáveis no tempo, embora possa ser estendido a certos sistemas de controle não lineares. (Referir-se ao Cap. 11).

Funções de transferência. A função de transferência de um sistema linear invariante no tempo é definida como sendo a relação da transformada de Laplace da saída (função resposta) para a transformada de Laplace da entrada (função excitação), considerando-se nulas todas as condições iniciais.

Considere o sistema linear invariável no tempo definido pela seguinte equação diferencial:

$$a_0 \overset{(n)}{y} + a_1 \overset{(n-1)}{y} + \cdots + a_{n-1}\dot{y} + a_n y = b_0 \overset{(m)}{x} + b_1 \overset{(m-1)}{x} + \cdots + b_{m-1}\dot{x} + b_m x \quad (n \geq m) \quad (4.1)$$

onde y é a saída do sistema e x é a entrada. A função de transferência deste sistema é obtida considerando-se as transformadas de Laplace de ambos os membros da Eq. (4.1), considerando-se que todas as condições iniciais são nulas, ou seja,

$$\text{Função de transferência} = G(s) = \frac{Y(s)}{X(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \cdots + b_{m-1}s + b_m}{a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}$$

A função de transferência é uma expressão relacionando a saída e a entrada de um sistema linear invariável no tempo em termos dos parâmetros do sistema e é uma propriedade do próprio sistema, independente da entrada ou função de excitação. A função de transferência inclui as unidades necessárias para relacionar a entrada com a saída; entretanto, não fornece qualquer informação relativa à estru-

tura física do sistema. (As funções de transferência de muitos sistemas físicos diferentes podem ser idênticas.)

Pela utilização deste conceito, pode-se representar a dinâmica do sistema por equações algébricas em s. A maior potência de s no denominador da função de transferência é igual à ordem do termo de maior derivada na saída. Se a maior potência de s é igual a n, o sistema é denominado um sistema de ordem n-ésima.

Sistema translacional mecânico. Considere o sistema amortecedor viscoso-mola-massa indicado na Fig. 4.2. Um amortecedor do tipo indicado é um dispositivo que proporciona uma fricção viscosa, ou amortecimento. Consiste em um pistão e um cilindro cheio de óleo. Qualquer movimento relativo entre a haste do pistão e o cilindro é resistido pelo óleo porque este deve fluir em volta do pistão (ou através de orifícios existentes no pistão) de um lado do pistão para o outro. O amortecedor essencialmente absorve energia. Esta energia absorvida é dissipada como calor, e o amortecedor não armazena qualquer energia cinética ou potencial.

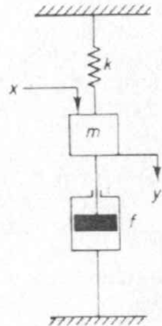


Fig. 4.2 Sistema do amortecedor viscoso-mola-massa.

Calculemos a função de transferência deste sistema admitindo que a força $x(t)$ é a entrada e o deslocamento da massa é a saída $y(t)$. Procederemos de acordo com os seguintes passos:

1. Escrever a equação diferencial do sistema.
2. Considerar a transformada de Laplace da equação diferencial, admitindo que todas as condições iniciais são nulas.
3. Calcular a relação entre a saída $Y(s)$ e a entrada $X(s)$. Esta relação é a função de transferência.

Para calcular a equação diferencial linear invariante no tempo, vamos supor que a força de fricção do amortecedor é proporcional a $\dot{y}$ e que a mola é uma mola linear, isto é, que a força na mola é proporcional a y. Neste sistema, m indica a massa, f indica o coeficiente da fricção-viscosa, e k denota a constante da mola.

A lei fundamental que governa os sistemas mecânicos é a lei de Newton. Para sistemas de translação a lei estabelece que

$$ma = \sum F$$

onde

m = massa, kg
a = aceleração, m/s²
F = força, N

Um quilograma é uma unidade de massa. Quando é acionado por uma força de 1 N, a massa de 1 kg acelera com 1 m/s².

Aplicando a lei de Newton ao sistema presente, obtemos

$$m\frac{d^2y}{dt^2} = -f\frac{dy}{dt} - ky + x$$

ou

$$m\frac{d^2y}{dt^2} + f\frac{dy}{dt} + ky = x \tag{4.2}$$

Considerando a transformada de Laplace de cada termo da Eq. (4.2), resulta

$$\mathscr{L}\left[m\frac{d^2y}{dt^2}\right] = m[s^2Y(s) - sy(0) - \dot{y}(0)]$$

$$\mathscr{L}\left[f\frac{dy}{dt}\right] = f[sY(s) - y(0)]$$

$$\mathscr{L}[ky] = kY(s)$$

$$\mathscr{L}[x] = X(s)$$

Se considerarmos as condições iniciais iguais a zero de modo que $y(0) = 0$, $\dot{y}(0) = 0$, a transformada de Laplace da Eq. (4.2) pode ser escrita

$$(ms^2 + fs + k)Y(s) = X(s)$$

Calculando a relação entre $Y(s)$ e $X(s)$, encontramos que a função de transferência do sistema é dada por

$$\text{Função de transferência} = G(s) = \frac{Y(s)}{X(s)} = \frac{1}{ms^2 + fs + k}$$

Sistema rotacional mecânico. Considere o sistema indicado na Fig. 4.3. O sistema consiste em uma carga de inércia e um amortecedor por fricção-viscosa.

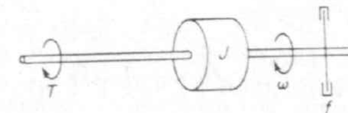


Fig. 4.3 Sistema rotacional mecânico.

Por definição,

J = momento de inércia da carga, kg-m^2
f = coeficiente de fricção-viscosa, N-m/rad/s
ω = velocidade angular, rad/s
T = torque aplicado ao sistema, N-m

Para sistemas rotacionais mecânicos, a lei de Newton estabelece que

$$J\alpha = \sum T$$

onde

J = momento de inércia, kg-m^2
α = aceleração angular, rad/s^2
T = torque, N-m

A relação entre g-cm^2 e slug-ft^2 é dada por

$$1 \text{ slug-ft}^2 = 13{,}56 \times 10^6 \text{ g-cm}^2$$

As unidades consistentes para massa, momento de inércia e torque são fornecidas pela seguinte tabela:

Massa	Momento de inércia	Torque
slug	slug-ft^2	N-m
grama	g-cm^2	dina-cm
quilograma	kg-m^2	newton-m

Aplicando a lei de Newton ao sistema apresentado, obtemos

$$J\dot{\omega} + f\omega = T$$

Se admitirmos que o torque T aplicado é a entrada e a velocidade angular ω é a saída, então a função de transferência deste sistema é dada por

$$\frac{\Omega(s)}{T(s)} = \frac{1}{Js + f}$$

onde

$$\Omega(s) = \mathscr{L}[\omega(t)]$$
$$T(s) = \mathscr{L}[T(t)]$$

Circuito L-R-C. Considere o circuito elétrico indicado na Fig. 4.4. O circuito consiste em uma indutância L (henry), uma resistência R (ohm), e uma capacitância

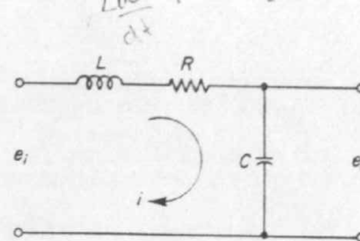


Fig. 4.4 Circuito elétrico.

C (farad). Aplicando as leis de Kirchhoff para o sistema, obtemos as seguintes equações:

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int i\,dt = e_i \tag{4.3}$$

$$\frac{1}{C}\int i\,dt = e_0 \tag{4.4}$$

Tomando as transformadas de Laplace das Eqs. (4.3) e (4.4), admitindo condições iniciais nulas, obtemos

$$LsI(s) + RI(s) + \frac{1}{C}\frac{1}{s}I(s) = E_i(s)$$

$$\frac{1}{C}\frac{1}{s}I(s) = E_0(s)$$

Se se admite que e_i é a entrada e e_0 a saída, então, demonstra-se que a função de transferência deste sistema é

$$\frac{E_0(s)}{E_i(s)} = \frac{1}{LCs^2 + RCs + 1} \tag{4.5}$$

Impedâncias complexas. Na dedução de funções de transferência para redes elétricas, freqüentemente verificamos ser conveniente escrever diretamente as equações transformadas em Laplace sem escrever as equações diferenciais. Considere o sistema indicado na Fig. 4.5(a). Neste sistema, Z_1 e Z_2 representam impedâncias complexas. A impedância complexa $Z(s)$ de um circuito com dois terminais é a relação entre $E(s)$, a transformada de Laplace da tensão entre os terminais, e

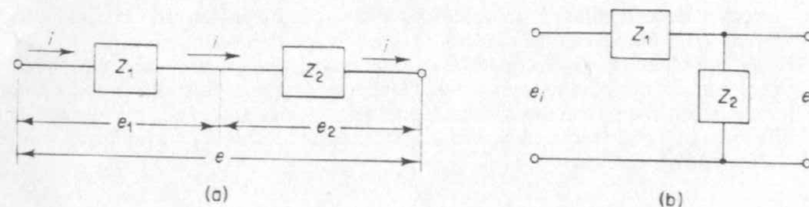


Fig. 4.5 Circuitos elétricos.

$I(s)$, a transformada de Laplace da corrente através do elemento, admitindo-se condições iniciais nulas, de modo que $Z(s) = E(s)/I(s)$. Se o elemento de dois terminais for uma resistência R, capacitância C ou indutância L, então a respectiva impedância complexa é dada por R, $1/Cs$, ou Ls, respectivamente. Se impedâncias complexas forem ligadas em série, a impedância total é a soma das impedâncias complexas individuais.

Considere o circuito indicado na Fig. 4.5(b). Suponha que as tensões e_i e e_0 correspondem à entrada e à saída do circuito, respectivamente. A função de transferência deste circuito é, então,

$$\frac{E_0(s)}{E_i(s)} = \frac{Z_2(s)}{Z_1(s) + Z_2(s)}$$

Para o sistema mostrado na Fig. 4.4,

$$Z_1 = Ls + R, \qquad Z_2 = \frac{1}{Cs}$$

Portanto, a função de transferência $E_0(s)/E_i(s)$ pode ser escrita:

$$\frac{E_0(s)}{E_i(s)} = \frac{\frac{1}{Cs}}{Ls + R + \frac{1}{Cs}}$$

$$= \frac{1}{LCs^2 + RCs + 1}$$

que, é claro, é idêntica à Eq. (4.5).

Elementos passivos e elementos ativos. Alguns dos elementos em um sistema (por exemplo, capacitâncias e indutâncias em um sistema elétrico) armazenam energia. Esta energia pode posteriormente ser introduzida no sistema. A quantidade de energia que pode ser introduzida não pode exceder a quantidade que o elemento armazenou, e, a menos que um elemento tenha armazenado energia anteriormente, ele não pode fornecer qualquer energia para um sistema. Por esta razão, tais elementos são denominados elementos *passivos*. Um sistema contendo apenas elementos passivos é denominado um sistema *passivo*. Exemplos de elementos passivos são as capacitâncias, resistências, indutâncias; massas, inércias, amortecedores, molas. Para sistemas passivos, todos os termos em uma equação diferencial homogênea do sistema possuem o mesmo sinal.

Um elemento físico que pode fornecer energia externa para um sistema é denominado um elemento *ativo*. Por exemplo, um amplificador é um elemento ativo desde que possui uma fonte de potência e supre potência ao sistema. Fontes de força externa, torque ou velocidade; fontes de tensão, ou corrente etc. também são elementos ativos.

Analogia força-tensão. Considere o sistema mecânico indicado na Fig. 4.6(a) e o sistema elétrico mostrado na Fig. 4.6(b). A equação diferencial para este sistema mecânico é

$$m\frac{d^2x}{dt^2} + f\frac{dx}{dt} + kx = p \tag{4.6}$$

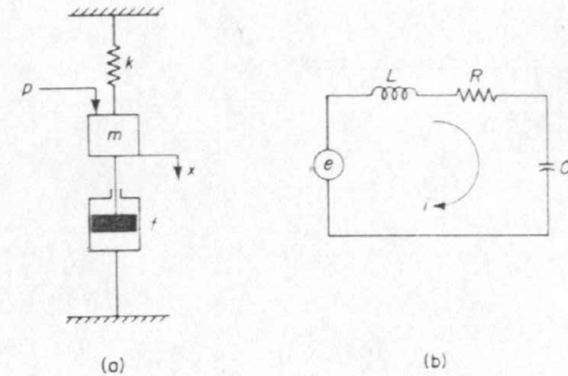


Fig. 4.6 (a) Sistema mecânico; (b) sistema elétrico análogo.

enquanto a equação diferencial para o sistema elétrico é

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int i\,dt = e$$

Em termos da carga elétrica q, esta última equação torna-se

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{1}{C}q = e \tag{4.7}$$

Comparando as Eqs. (4.6) e (4.7), verificamos que as equações diferenciais para os dois sistemas são idênticas. Estes sistemas são denominados sistemas análogos, e os termos que ocupam posições correspondentes nas equações diferenciais são chamados de grandezas análogas. Uma lista de grandezas análogas é fornecida na Tabela 4.1. Neste caso a analogia é denominada analogia força-tensão elétrica.

Tabela 4.1 Grandezas análogas na analogia força-tensão

Sistema mecânico	Sistema elétrico
força p (torque T)	tensão elétrica e
massa m (momento de inércia J)	indutância L
coeficiente de fricção-viscosa f	resistência R
constante da mola k	recíproca de capacitância (elastância) $1/C$
deslocamento x (deslocamento angular θ)	carga q
velocidade $\dot{x}$ (velocidade angular $\dot{\theta}$)	corrente i

A Fig. 4.7 mostra alguns exemplos de sistemas análogos. Cada um dos sistemas elétricos e seu sistema mecânico correspondente possuem funções de transferência análogas. (Na Fig. 4.7, x_i e x_0 indicam deslocamentos.) Note que na dedução das funções de transferência, consideramos que os sistemas são constituídos por parâmetros concentrados e que não há efeito de carga na saída.

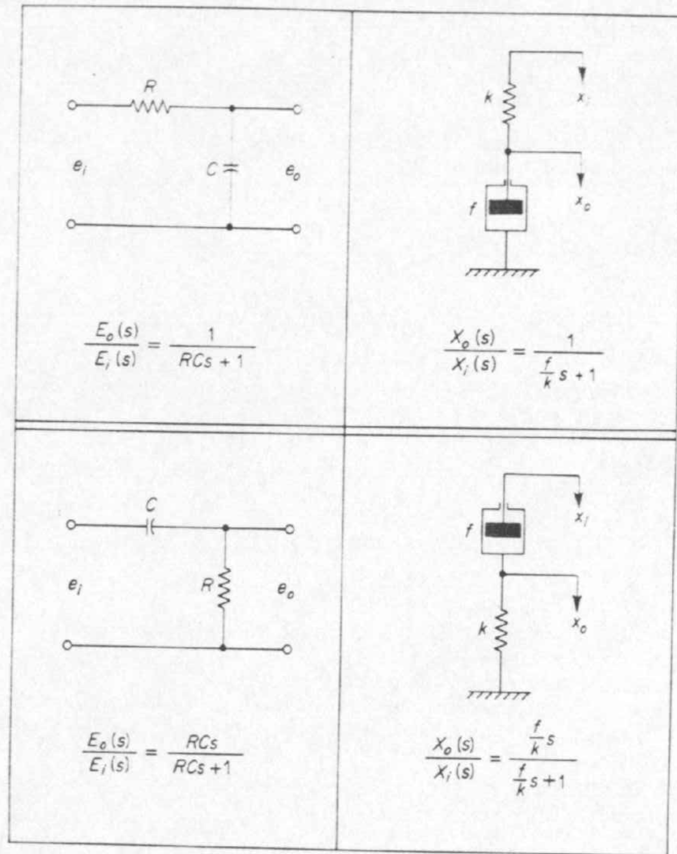


Fig. 4.7 Sistemas análogos.

Analogia força-corrente. Outra analogia útil entre sistemas elétricos e sistemas mecânicos é baseada na analogia força-corrente. Considere o sistema mecânico mostrado na Fig. 4.8(a). A equação diferencial que descreve o sistema é

$$m\frac{d^2x}{dt^2} + f\frac{dx}{dt} + kx = p \tag{4.8}$$

Considere agora o sistema elétrico mostrado na Fig. 4.8(b). Aplicando a lei de Kirchhoff relativa a correntes, obtemos

$$i_L + i_R + i_C = i_s \tag{4.9}$$

onde

$$i_L = \frac{1}{L}\int e\,dt$$

$$i_R = \frac{e}{R}$$

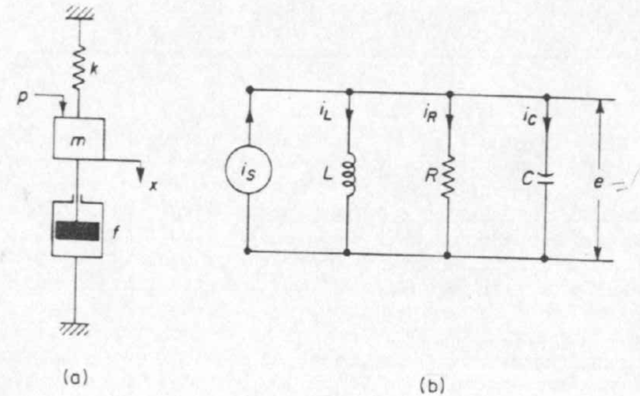


Fig. 4.8 (a) Sistema mecânico; (b) sistema elétrico análogo.

$$i_C = C\frac{de}{dt}$$

A Eq. (4.9) pode ser escrita

$$\frac{1}{L}\int e\,dt + \frac{e}{R} + C\frac{de}{dt} = i_s \tag{4.10}$$

Note que o fluxo magnético concatenado ψ é relacionado com e pela seguinte equação:

$$\frac{d\psi}{dt} = e$$

Em termos de ψ, a Eq. (4.10) pode ser escrita como segue:

$$C\frac{d^2\psi}{dt^2} + \frac{1}{R}\frac{d\psi}{dt} + \frac{1}{L}\psi = i_s \tag{4.11}$$

Comparando as Eqs. (4.8) e (4.11), verificamos que os dois sistemas são sistemas análogos. As grandezas análogas são listadas na Tabela 4.2. Neste caso a analogia é denominada analogia força-corrente.

Tabela 4.2 Grandezas análogas na analogia força-corrente

Sistema mecânico	Sistema elétrico
força p (torque T)	corrente elétrica i
massa m (inércia J)	capacitância C
coeficiente de fricção-viscosa f	recíproco da resistência (condutância) $1/R$
constante da mola k	recíproco da indutância $1/L$
deslocamento x (deslocamento angular θ)	fluxo magnético concatenado ψ
velocidade $\dot{x}$ (velocidade angular $\dot{\theta}$)	tensão elétrica e

Sistemas análogos. O conceito de sistemas análogos é muito útil na prática desde que um tipo de sistema pode ser mais fácil de ser manuseado experimentalmente do que um outro tipo. Por exemplo, em vez de construir e estudar um sistema mecânico, podemos construir e estudar seu análogo elétrico porque, em geral, sistemas elétricos ou eletrônicos são muito mais fáceis de serem analisados experimentalmente. Em particular, computadores analógicos eletrônicos são muito úteis para simulação de sistemas mecânicos, bem como de outros sistemas físicos.

Deve ser lembrado que analogias entre estes sistemas caem por terra quando as regiões de operação são muito extensas. Em outras palavras, desde que as equações diferenciais relacionadas com as analogias são baseadas somente por aproximação às características dinâmicas dos sistemas físicos, a analogia pode falhar se a região de operação de um sistema é muito ampla. Se a região de operação de um dado sistema mecânico é ampla, entretanto, pode ser dividida em duas ou mais sub-regiões e os sistemas elétricos análogos podem ser elaborados para cada sub-região. De fato, analogias não são limitadas a sistemas elétricos e sistemas mecânicos; são aplicáveis a quaisquer sistemas nos quais suas equações diferenciais, ou funções de transferência, são de forma idêntica.

Funções de transferência de elementos em cascata. Muitos sistemas realimentados possuem componentes que carregam uns aos outros. Considere o sistema mostrado na Fig. 4.9. Suponha que e_i é a entrada e e_0 é a saída. Neste sistema o

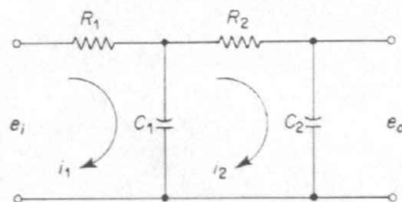


Fig. 4.9 Sistema elétrico.

segundo estágio do circuito (parte R_2C_2) produz um efeito de carga no primeiro estágio (parte R_1C_1). As equações para este sistema são

$$\frac{1}{C_1}\int (i_1 - i_2)\,dt + R_1 i_1 = e_i \tag{4.12}$$

e

$$\frac{1}{C_1}\int (i_2 - i_1)\,dt + R_2 i_2 = -\frac{1}{C_2}\int i_2\,dt = -e_0 \tag{4.13}$$

Tomando as transformadas de Laplace das Eqs. (4.12) e (4.13), admitindo condições iniciais nulas, obtemos

$$\frac{1}{C_1 s}[I_1(s) - I_2(s)] + R_1 I_1(s) = E_i(s) \tag{4.14}$$

$$\frac{1}{C_1 s}[I_2(s) - I_1(s)] + R_2 I_2(s) = -\frac{1}{C_2 s} I_2(s) = -E_0(s) \tag{4.15}$$

Eliminando $I_1(s)$ e $I_2(s)$ das Eqs. (4.14) e (4.15), verificamos que a função de transferência entre $E_0(s)$ e $E_i(s)$ é

$$\begin{aligned}\frac{E_0(s)}{E_i(s)} &= \frac{1}{(R_1C_1s + 1)(R_2C_2s + 1) + R_1C_2s} \\ &= \frac{1}{R_1C_1R_2C_2s^2 + (R_1C_1 + R_2C_2 + R_1C_2)s + 1}\end{aligned} \tag{4.16}$$

O termo R_1C_2s no denominador da função de transferência representa a interação dos dois circuitos simples RC. Desde que $(R_1C_1 + R_2C_2 + R_1C_2)^2 > 4R_1C_1R_2C_2$, as duas raízes do denominador da Eq. (4.16) são reais.

Esta análise demonstrou que se dois circuitos RC são ligados em cascata, de modo que a saída do primeiro circuito é a entrada do segundo, a função de transferência global não é o produto de $1/(R_1C_1s + 1)$ e $1/(R_2C_2s + 1)$. A justificativa para este fato é que, quando calculamos a função de transferência para um circuito isolado, admitimos implicitamente que a saída não é carregada. Em outras palavras, se admite a impedância de carga infinita, o que significa que nenhuma potência está sendo consumida na saída. Quando o segundo circuito é ligado na saída do primeiro, entretanto, uma certa quantidade de potência é consumida, sendo conseqüentemente violada a consideração de não haver carga na saída. Portanto, se a função de transferência deste sistema é obtida sobre a hipótese de não haver carga, então ela não é válida. O grau de efeito de carregamento determina a quantidade de modificação da função de transferência.

Funções de transferência de elementos em cascata sem carregamento. A função de transferência de um sistema consistindo em dois elementos em cascata sem carregamento pode ser obtida eliminando-se a entrada e a saída intermediárias. Por exemplo, considere o sistema mostrado na Fig. 4.10(a). A função de transferência de cada elemento é

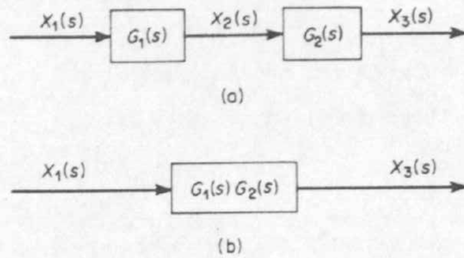


Fig. 4.10 (a) Sistema consistindo em dois elementos em cascata sem carregamento; (b) um sistema equivalente.

$$G_1(s) = \frac{X_2(s)}{X_1(s)}, \qquad G_2(s) = \frac{X_3(s)}{X_2(s)}$$

Se a impedância de entrada do segundo elemento é infinita, a saída do primeiro elemento não é afetada pela conexão ao segundo elemento. Neste caso, a função de transferência de todo o sistema é

$$G(s) = \frac{X_3(s)}{X_1(s)} = \frac{X_2(s)}{X_1(s)}\frac{X_3(s)}{X_2(s)} = G_1(s)G_2(s)$$

A função de transferência do sistema global é então o produto da função de transferência dos elementos individuais. Isto é mostrado na Fig. 4.10(b).

Como exemplo, considere o sistema indicado na Fig. 4.11. A inserção de um amplificador de isolação entre os circuitos para obter características sem carregamento é freqüentemente utilizada para combinar circuitos elétricos. Desde que tanto amplificadores de estado sólido como amplificadores a válvulas possuem impedâncias de entrada muito altas, um amplificador de isolação inserido entre os dois circuitos justifica a afirmação de não haver carregamento.

Os dois circuitos simples RC, isolados por um amplificador conforme indicado na Fig. 4.11, possui efeitos de carregamento desprezíveis e a função de transferência de todo o circuito é igual ao produto das funções de transferência individuais. Conseqüentemente, neste caso,

$$\frac{E_0(s)}{E_i(s)} = \left(\frac{1}{R_1C_1s+1}\right)(K)\left(\frac{1}{R_2C_2s+1}\right)$$
$$= \frac{K}{(R_1C_1s+1)(R_2C_2s+1)}$$



Fig. 4.11 Sistema elétrico.

4.3 LINEARIZAÇÃO DE UM MODELO MATEMÁTICO NÃO LINEAR

Esta seção apresenta uma técnica de linearização aplicável a muitos sistemas não lineares. Aplicaremos esta técnica a um servomotor hidráulico e obteremos uma função de transferência para o servomotor hidráulico linearizado.

Aproximação linear de sistemas não lineares. A fim de obter um modelo matemático linear para um sistema não linear, suporemos que as variáveis variam muito pouco em relação a alguma condição de operação. Considere um sistema cuja entrada é $x(t)$ e cuja saída é $y(t)$. A relação entre $y(t)$ e $x(t)$ é dada por

$$y = f(x) \tag{4.17}$$

Se a condição de operação normal corresponde a $\bar{x}, \bar{y}$, então a Eq. (4.17) pode ser expandida em série de Taylor em torno desse ponto como segue:

$$y = f(x)$$
$$= f(\bar{x}) + \frac{df}{dx}(x-\bar{x}) + \frac{1}{2!}\frac{d^2f}{dx^2}(x-\bar{x})^2 + \cdots \tag{4.18}$$

onde as derivadas df/dx, d^2f/dx^2,... são calculadas em $x = \bar{x}$. Se a variação de $x - \bar{x}$ é pequena, podemos desprezar os termos de maior ordem em $x - \bar{x}$. Então a Eq. (4.18) pode ser escrita

$$y = \bar{y} + K(x-\bar{x}) \tag{4.19}$$

onde

$$\bar{y} = f(\bar{x})$$
$$K = \left.\frac{df}{dx}\right|_{x=\bar{x}}$$

A Eq. (4.19) pode ser reescrita como

$$y - \bar{y} = K(x-\bar{x}) \tag{4.20}$$

o que indica que $y - \bar{y}$ é proporcional a $x - \bar{x}$. A Eq. (4.20) fornece um modelo matemático linear para um sistema não linear dado pela Eq. (4.17).

Agora considere um sistema não linear cuja saída y é uma função de duas entradas x_1 e x_2, de modo que

$$y = f(x_1, x_2) \tag{4.21}$$

De modo a obter uma aproximação linear para este sistema não linear, podemos expandir a Eq. (4.21) em uma série de Taylor em torno do ponto de operação $\bar{x}_1, \bar{x}_2$. Então a Eq. (4.21) torna-se

$$y = f(\bar{x}_1, \bar{x}_2) + \left[\frac{\partial f}{\partial x_1}(x_1 - \bar{x}_1) + \frac{\partial f}{\partial x_2}(x_2 - \bar{x}_2)\right]$$
$$+ \frac{1}{2!}\left[\frac{\partial^2 f}{\partial x_1^2}(x_1 - \bar{x}_1)^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)\right.$$
$$\left. + \frac{\partial^2 f}{\partial x_2^2}(x_2 - \bar{x}_2)^2\right] + \cdots$$

onde as derivadas parciais são calculadas em $x_1 = \bar{x}_1, x_2 = \bar{x}_2$. Próximo ao ponto de operação normal, os termos de ordem superior podem ser desprezados. O modelo matemático linear deste sistema não linear na vizinhança da condição da operação normal é então dado por

$$y - \bar{y} = K_1(x_1 - \bar{x}_1) + K_2(x_2 - \bar{x}_2)$$

onde

$$\bar{y} = f(\bar{x}_1, \bar{x}_2)$$
$$K_1 = \left.\frac{\partial f}{\partial x_1}\right|_{x_1 = \bar{x}_1,\, x_2 = \bar{x}_2}$$
$$K_2 = \left.\frac{\partial f}{\partial x_2}\right|_{x_1 = \bar{x}_1,\, x_2 = \bar{x}_2}$$

Servomotor hidráulico. A Fig. 4.12 mostra um servomotor hidráulico. É essencialmente um amplificador de potência hidráulico controlado por válvula piloto e um atuador. A válvula piloto é uma válvula balanceada no sentido de que as forças de pressão que agem na mesma são todas balanceadas. Uma potência de saída

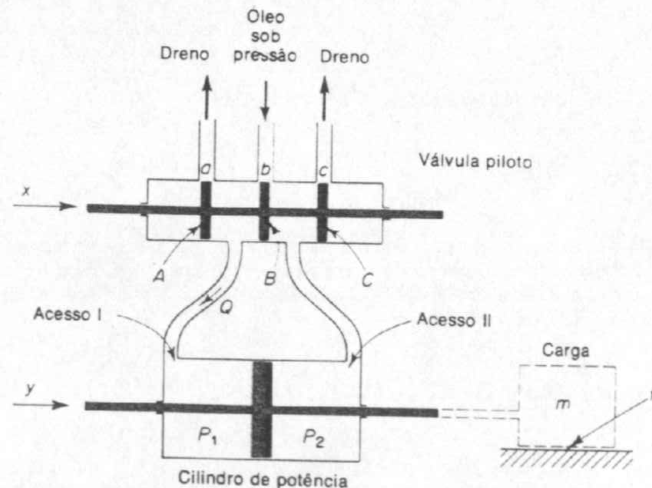


Fig. 4.12 Diagrama esquemático de um servomotor hidráulico.

muito grande pode ser controlada por uma válvula piloto, a qual pode ser posicionada com uma potência muito pequena.

A operação do servomotor hidráulico é a seguinte: Se a válvula piloto é movida para a direita, então a abertura I é ligada à abertura de suprimento e óleo sob pressão entra no lado esquerdo do pistão de potência. Desde que a abertura II é ligada à abertura de dreno, o óleo no lado direito do pistão de potência é retornado ao dreno. O óleo fluindo no cilindro de potência está sob alta pressão e o óleo fluindo para fora do cilindro de potência através do dreno está sob baixa pressão. A diferença de pressão resultante em ambos os lados do pistão de potência ocasiona um movimento para a direita. O óleo retornado é pressurizado por uma bomba e recirculado no sistema. Quando o pistão piloto é movido para a esquerda, o pistão de potência move-se para a esquerda.

Na prática, as aberturas *a*, *b* e *c* mostradas na Fig. 4.12 são sempre construídas maiores do que as válvulas correspondentes *A*, *B* e *C*. Neste caso há sempre uma fuga através da válvula. Este fato aumenta tanto a sensibilidade como a linearidade do servomotor hidráulico. Consideraremos esta hipótese na análise que se segue. [Note que algumas vezes um sinal "trêmulo", um sinal de alta freqüência de amplitude muito pequena (em relação ao deslocamento máximo da válvula), é sobreposto ao movimento da válvula piloto. Isto também aumenta a sensibilidade e a linearidade. Neste caso também há fuga através da válvula.]

Vamos definir

Q = taxa de fluxo de óleo para o cilindro de potência, kg/s
$\Delta P = P_2 - P_1$ = diferença de pressão através do pistão de potência, N/m²
x = deslocamento da válvula piloto, m.

Na Fig. 4.12, pode-se ver que Q é uma função de x e ΔP. Em geral, a relação entre as variáveis Q, x e ΔP é dada por uma equação não linear:

$$Q = f(x, \Delta P)$$

Linearizando esta equação não linear em torno do ponto de operação normal $\bar{Q}$, $\bar{x}$, $\Delta\bar{P}$, obtemos

$$Q - \bar{Q} = K_1(x - \bar{x}) - K_2(\Delta P - \Delta\bar{P}) \tag{4.22}$$

onde

$$\bar{Q} = f(\bar{x}, \Delta\bar{P})$$
$$K_1 = \left.\frac{\partial Q}{\partial x}\right|_{x = \bar{x},\, \Delta P = \Delta\bar{P}}$$
$$K_2 = -\left.\frac{\partial Q}{\partial\, \Delta P}\right|_{x = \bar{x},\, \Delta P = \Delta\bar{P}}$$

Note que para este sistema, a condição de operação normal corresponde a $\bar{Q} = 0$, $\bar{x} = 0$ e $\Delta\bar{P} = 0$. Portanto, obtemos da Eq. (4.22),

$$Q = K_1 x - K_2\, \Delta P \tag{4.23}$$

A Fig. 4.13 mostra esta relação linearizada entre Q, x e ΔP. As retas indicadas correspondem às curvas características do servomotor hidráulico linearizado. Esta família de curvas consiste em retas paralelas eqüidistantes, parametrizadas por x.

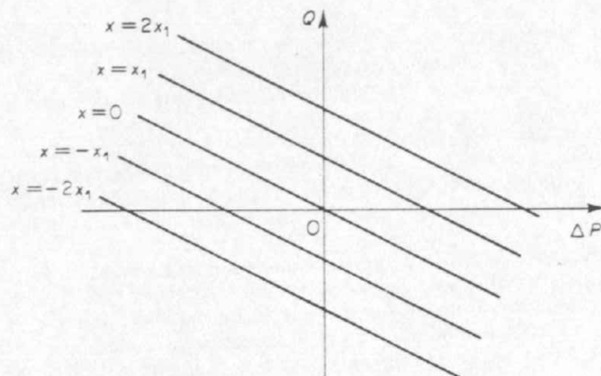


Fig. 4.13 Curvas características do servomotor hidráulico linearizado.

Referindo-se à Fig. 4.12, notamos que a taxa de fluxo de óleo Q (kg/s) vezes dt (s) é igual ao deslocamento do pistão de potência dy (m) vezes a área do pistão A (m²) vezes a densidade do óleo ρ (kg/m³). Portanto, obtemos

$$A\rho\, dy = Q\, dt$$

Note que para uma dada taxa de fluxo Q, quanto maior for a área do pistão A menor será a velocidade dy/dt. Portanto, se a área do pistão A é feita menor, permanecendo as outras variáveis constantes, a velocidade dy/dt tornar-se-á maior. Também uma taxa de fluxo Q maior acarretará um aumento na velocidade do pistão de potência e um menor tempo de resposta.

A Eq. (4.23) pode agora ser escrita como

$$\Delta P = \frac{1}{K_2}\left(K_1 x - A\rho\frac{dy}{dt}\right)$$

A força desenvolvida pelo pistão de potência é igual à diferença de pressão ΔP vezes a área do pistão A ou

$$\text{Força desenvolvida pelo pistão de potência} = A\ \Delta P$$

$$= \frac{A}{K_2}\left(K_1 x - A\rho\frac{dy}{dt}\right)$$

Para uma dada força máxima, se a diferença de pressão for suficientemente alta, a área do pistão ou o volume do óleo no cilindro pode ser feito pequeno. Conseqüentemente, para minimizar o peso do controlador, devemos fazer a pressão de suprimento suficientemente alta.

Suponha que o pistão de potência mova uma carga que consiste em uma massa e uma fricção-viscosa. Então, a força desenvolvida pelo pistão de potência é aplicada à massa de carga e à fricção, obtendo-se

$$m\ddot{y} + f\dot{y} = \frac{A}{K_2}(K_1 x - A\rho\dot{y})$$

ou

$$m\ddot{y} + \left(f + \frac{A^2\rho}{K_2}\right)\dot{y} = \frac{AK_1}{K_2}x \qquad (4.24)$$

onde m é a massa (kg) da carga e f é o coeficiente de fricção viscosa (N-s/m).

Supondo que o deslocamento da válvula piloto x é a entrada e o deslocamento do pistão de potência y a saída, verificamos que a função de transferência para o servomotor hidráulico é, a partir da Eq. (4.24),

$$\frac{Y(s)}{X(s)} = \frac{1}{s\left[\left(\frac{mK_2}{AK_1}\right)s + \frac{fK_2}{AK_1} + \frac{A\rho}{K_1}\right]}$$

$$= \frac{K}{s(Ts+1)} \qquad (4.25)$$

onde

$$K = \frac{1}{\frac{fK_2}{AK_1} + \frac{A\rho}{K_1}} \qquad \text{e} \qquad T = \frac{mK_2}{fK_2 - A^2\rho}$$

Da Eq. (4.25), pode-se notar que esta função de transferência é de segunda ordem. Se a relação $mK_2/(fK_2 + A^2\rho)$ é suficientemente pequena a ponto de poder ser desprezada ou a constante de tempo T for desprezível, a função de transferência pode ser simplificada para resultar

$$\frac{Y(s)}{X(s)} = \frac{K}{s}$$

Uma análise mais detalhada indica que, se a fuga do óleo, compressibilidade (incluindo os efeitos do ar dissolvido), expansão de tubulação etc. são levados em consideração, a função de transferência torna-se

$$\frac{Y(s)}{X(s)} = \frac{K}{s(T_1 s + 1)(T_2 s + 1)}$$

onde T_1 e T_2 são constantes de tempo. De fato, estas constantes de tempo dependem do volume de óleo do circuito de operação. Quanto menor for o volume, menores serão as constantes de tempo.

4.4 DIAGRAMAS DE BLOCOS

Um sistema de controle pode consistir em vários componentes. A fim de mostrar as funções desempenhadas por componentes, em engenharia de controle usa-se quase sempre um diagrama denominado "diagrama de blocos".

Diagrama de blocos. Um diagrama de blocos de um sistema é uma representação das funções desempenhadas por cada componente e do fluxo de sinais. Este diagrama indica a inter-relação que existe entre os vários componentes. Diferentemente de uma representação matemática, puramente abstrata, um diagrama de blocos possui a vantagem de indicar mais realisticamente os fluxos de sinal do sistema real.

Em um diagrama de blocos, todas as variáveis do sistema são ligadas às outras através de blocos funcionais. O "bloco funcional", ou simplesmente "bloco" é um símbolo para a operação matemática sobre o sinal de entrada para o bloco que produz a saída. As funções de transferência dos componentes são usualmente indicadas nos blocos correspondentes, os quais são ligados por flechas para indicar o sentido do fluxo de sinais. Note que o sinal pode passar somente no sentido da flecha. Conseqüentemente, um diagrama de blocos de um sistema de controle indica explicitamente uma propriedade unilateral.

A Fig. 4.14 mostra um elemento de diagrama de blocos. O segmento orientado (flecha) apontando para o bloco indica a entrada e o segmento orientado saindo do bloco representa a saída. Estas flechas são referidas como sinais.

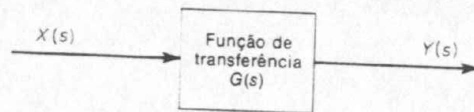


Fig. 4.14 Elemento de um diagrama de bloco.

Note que as dimensões do sinal de saída do bloco são as dimensões do sinal de entrada multiplicado pelas dimensões da função de transferência do bloco.

As vantagens da representação em diagrama de blocos de um sistema são devidas ao fato de que é mais fácil construir o diagrama de bloco total para todo o sistema simplesmente conectando os blocos dos componentes de acordo com o fluxo de sinal, sendo também possível calcular a contribuição de cada componente no desempenho global do sistema.

Em geral, a operação funcional do sistema pode ser visualizada mais diretamente pelo exame do diagrama de blocos do que pelo exame do próprio sistema físico. Um diagrama de bloco contém informação relativa ao comportamento dinâmico, porém não contém qualquer informação relativa à construção física do sistema. Portanto, muitos sistemas completamente diferentes e sem nenhuma relação podem ser representados pelo mesmo diagrama de blocos.

Deve ser notado que em um diagrama de blocos a principal fonte de energia não é explicitamente mostrada e, também, que um diagrama de blocos de um dado sistema não é único. Vários diagramas de blocos diferentes podem ser desenhados para um dado sistema, dependendo dos objetivos da análise.

Detector de erro. O detector de erro produz um sinal que é a diferença entre a entrada de referência e o sinal realimentado do sistema de controle. No projeto, a

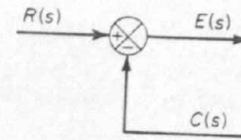


Fig. 4.15 Diagrama de bloco de um detector de erro.

escolha do detector de erro é muito importante e deve ser decidida cuidadosamente. Isto é devido ao fato de que quaisquer imperfeições no detector de erros inevitavelmente prejudicaria o desempenho de todo o sistema. O diagrama de bloco de um detector de erro é indicado na Fig. 4.15.

Note que um círculo com uma cruz é o símbolo que indica uma operação de soma. Os sinais mais ou menos em cada flecha indicam se aquele sinal deve ser adicionado ou subtraído. É importante que as grandezas a serem adicionadas ou subtraídas possuam as mesmas dimensões e as mesmas unidades.

Diagrama de blocos de um sistema de malha-fechada. A Fig. 4.16 mostra um exemplo de um diagrama de blocos de um sistema de malha-fechada. A saída $C(s)$ é realimentada ao ponto de soma, onde é comparada com a entrada de referência $R(s)$. A natureza de malha-fechada do sistema é indicada claramente pela figura. A saída do bloco, $C(s)$ neste caso, é obtida pela multiplicação da função de transferência $G(s)$ com a entrada do bloco, $E(s)$.

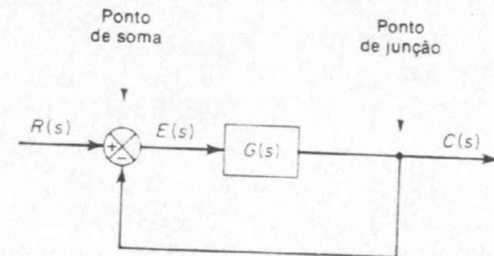


Fig. 4.16 Diagrama de bloco de um sistema em malha-fechada.

Qualquer sistema de controle linear pode ser representado por um diagrama de blocos que consiste em blocos, pontos de soma e pontos de junção. Um ponto de junção é um ponto do qual o sinal de saída de um bloco é ligado a vários outros blocos concomitantemente ou a pontos de soma.

Quando a saída é realimentada ao ponto de soma para comparação com a entrada é necessário converter a forma do sinal de saída para aquela do sinal de entrada. Por exemplo, em um sistema de controle de temperatura, o sinal de saída é usualmente a temperatura controlada. O sinal de saída, que possui a dimensão de temperatura, deve ser convertido para uma força ou posição antes de ser comparado com o sinal de entrada. Esta conversão é efetuada pelo elemento de realimentação cuja função de transferência é $H(s)$, como indicado na Fig. 4.17. Um outro papel importante do elemento de realimentação é modificar a saída antes de ser comparada com a entrada. No exemplo presente, o sinal realimentado que é enviado para trás até o ponto de soma para comparação com a entrada é $B(s) = H(s)C(s)$.

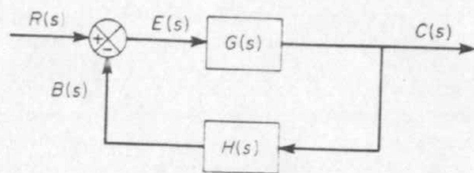


Fig. 4.17 Sistema em malha-fechada.

A relação entre o sinal realimentado $B(s)$ e o sinal erro atuante $E(s)$ é denominada função de transferência de malha-aberta. Isto é,

$$\text{Função de transferência de malha-aberta} = \frac{B(s)}{E(s)} = G(s)H(s)$$

A relação entre a saída $C(s)$ e o sinal erro atuante $E(s)$ é denominada função de transferência do ramo direto, de modo que,

$$\text{Função de transferência do ramo direto} = \frac{C(s)}{E(s)} = G(s)$$

Se a função de transferência realimentada é unitária, então a função de transferência de malha-aberta e a função de transferência do ramo direto são as mesmas. Para o sistema mostrado na Fig. 4.17, a saída $C(s)$ e a entrada $R(s)$ são relacionadas como segue:

$$C(s) = G(s)E(s)$$
$$E(s) = R(s) - B(s)$$
$$= R(s) - H(s)C(s)$$

Eliminando $E(s)$ destas equações, resulta

$$C(s) = G(s)[R(s) - H(s)C(s)]$$

ou

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} \tag{4.26}$$

A função de transferência relacionando $C(s)$ com $R(s)$ é a chamada função de transferência de malha-fechada. Esta função de transferência relaciona a dinâmica do sistema em malha-fechada para a dinâmica dos elementos do ramo direto e os elementos da realimentação.

Da Eq. (4.26), $C(s)$ é dada por

$$C(s) = \frac{G(s)}{1 + G(s)H(s)} R(s)$$

Portanto a saída do sistema de malha-fechada depende, obviamente, tanto da função de transferência de malha-fechada como da natureza da entrada.

Sistema em malha-fechada sujeito a um distúrbio. A Fig. 4.18 mostra um sistema em malha-fechada sujeito a um distúrbio. Quando duas entradas (a entrada de referência e o distúrbio) estão presentes em um sistema linear, cada entrada pode ser considerada independentemente da outra; e as saídas correspondentes para cada entrada sozinha podem ser adicionadas para resultar a saída completa. O modo pelo qual cada entrada é introduzida no sistema é indicado no ponto de soma ou por um sinal mais ou por um sinal menos.

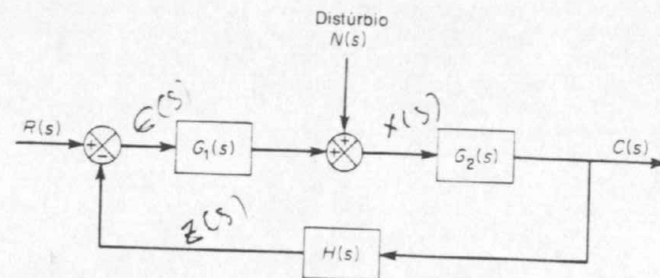


Fig. 4.18 Sistema em malha-fechada sujeito a um distúrbio.

Considere o sistema mostrado na Fig. 4.18. Examinando o efeito do distúrbio $N(s)$, podemos admitir que o sistema está em repouso inicialmente com erro zero; podemos então calcular a resposta $C_N(s)$ apenas devido ao distúrbio. Esta resposta é dada por

$$\frac{C_N(s)}{N(s)} = \frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)}$$

Por outro lado, considerando a resposta em relação à entrada de referência $R(s)$, podemos supor que o distúrbio é nulo. Então, a resposta $C_R(s)$ em relação à entrada de referência $R(s)$ pode ser obtida de

$$\frac{C_R(s)}{R(s)} = \frac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)H(s)}$$

A resposta com a aplicação simultânea da entrada de referência e do distúrbio pode ser obtida adicionando-se as duas respostas individuais. Em outras palavras, a resposta $C(s)$ devida à aplicação simultânea da entrada de referência $R(s)$ e ao distúrbio $N(s)$ é dada por

$$C(s) = C_R(s) + C_N(s)$$
$$= \frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)}[G_1(s)R(s) + N(s)]$$

Considere agora o caso onde $|G_1(s)H(s)| >> 1$ e $|G_1(s)G_2(s)H(s)| >> 1$. Neste caso, a função de transferência em malha-fechada $C_N(s)/N(s)$ torna-se quase nula e o efeito do distúrbio é eliminado. Esta é uma vantagem do sistema em malha-fechada.

Por outro lado, a função de transferência em malha-fechada $C_R(s)/R(s)$ aproxima-se a $1/H(s)$ conforme o ganho de $G_1(s)G_2(s)H(s)$ aumenta. Isto significa que se $|G_1(s)G_2(s)H(s)| >> 1$, então a função de transferência em malha-fechada $C_R(s)/R(s)$ torna-se independente de $G_1(s)$ e $G_2(s)$ e inversamente proporcional a $H(s)$, de modo que as variações de $G_1(s)$ e $G_2(s)$ não afetam a função de transferência em malha-fechada $C_R(s)/R(s)$. Esta é outra vantagem do sistema em malha-fechada. Pode ser facilmente verificado que qualquer sistema em malha-fechada com realimentação unitária, $H(s) = 1$, tende a igualar a entrada e a saída.

Procedimentos para construir diagramas de blocos. Para desenhar o diagrama de blocos de um sistema, escreva primeiro as equações que descrevem o comportamento dinâmico de cada componente, tome então as transformadas de Laplace das equações, admitindo condições iniciais nulas, e represente cada equação transformada de Laplace individualmente na forma de bloco. Finalmente, monte os elementos em um diagrama de blocos completo.

Como um exemplo, considere o circuito RC indicado na Fig. 4.19(a). As equações para este circuito são:

$$i = \frac{e_i - e_0}{R} \tag{4.27}$$

$$e_0 = \frac{\int i\,dt}{C} \tag{4.28}$$

As transformadas de Laplace das Eqs. (4.27) e (4.28), com uma condição inicial nula, tornam-se

$$I(s) = \frac{E_i(s) - E_0(s)}{R} \tag{4.29}$$

$$E_0(s) = \frac{I(s)}{Cs} \tag{4.30}$$

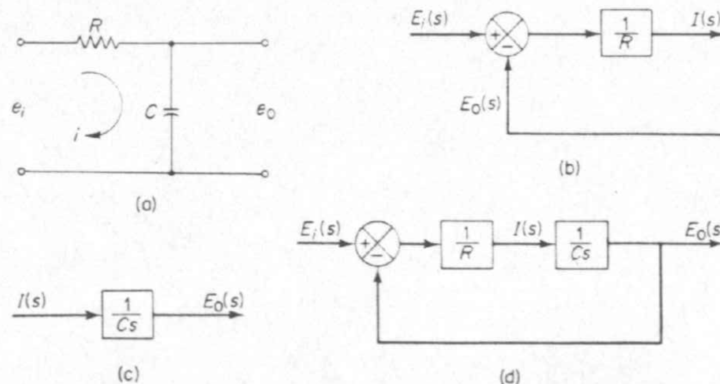


Fig. 4.19 (a) Circuito RC; (b) diagrama de blocos representando a Eq. (4.29); (c) diagrama de blocos representando a Eq. (4.30); (d) diagrama de blocos do circuito RC.

A Eq. (4.29) representa uma operação de soma e o diagrama correspondente é mostrado na Fig. 4.19(b). A Eq. (4.30) representa o bloco indicado na Fig. 4.19(c). Montando estes dois elementos, obtemos o diagrama de blocos global para o sistema, conforme indicado na Fig. 4.19(d).

Redução de diagramas de bloco. É importante notar que blocos podem ser conectados em série somente se a saída de um bloco não é afetada pelo bloco seguinte. Se houver qualquer efeito de carregamento entre os componentes, é necessário combinar estes componentes em um bloco simples. Em outras palavras, um único bloco representará as duas partes.

Tabela 4.3 Regras da álgebra de diagramas de blocos

	Diagramas de blocos originais	Diagramas de blocos equivalentes
1	A, B, C; A−B; A−B+C	A, C, B; A+C; A−B+C
2	A, B, C; A−B+C	A, B, C; A−B; A−B+C
3	A; G_1; AG_1; G_2; AG_1G_2	A; G_2; AG_2; G_1; AG_1G_2
4	A; G_1; AG_1; G_2; AG_1G_2	A; G_1G_2; AG_1G_2
5	A; G_1; AG_1; G_2; AG_2; AG_1+AG_2	A; G_1+G_2; AG_1+AG_2
6	A; G; AG; B; AG−B	A; $\frac{B}{G}$; $A-\frac{B}{G}$; G; AG−B; $\frac{1}{G}$; B
7	A; B; A−B; G; AG−BG	A; G; AG; B; G; BG; AG−BG

	Diagramas de blocos originais	Diagramas de blocos equivalentes
8	A, G, AG, AG	A, G, AG, G, AG
9	A, G, AG, A	A, G, AG, AG, $\frac{1}{G}$, A
10	A, B, A−B, A−B	B, A, A−B, B, A−B
11	A, G_1, AG_1, AG_1+AG_2, G_2, AG_2	A, G_2, AG_2, $\frac{1}{G_2}$, A, G_1, AG_1, AG_1+AG_2
12	A, G_1, B, G_2	A, $\frac{1}{G_2}$, G_2, G_1, B
13	A, G_1, B, G_2	A, $\frac{G_1}{1+G_1G_2}$, B

Qualquer número de blocos em cascata representando componentes sem efeito de carga podem ser substituídos por um único bloco cuja função de transferência é simplesmente o produto das funções de transferência individuais.

Um diagrama de blocos complicado envolvendo muitos laços de realimentação pode ser simplificado por um rearranjo passo a passo, usando regras de álgebra de diagrama de blocos. Algumas destas importantes regras são fornecidas na Tabela 4.3. Elas são obtidas escrevendo-se a mesma equação de uma maneira diferente. A simplificação do diagrama de blocos por rearranjos e substituições reduz consideravelmente o trabalho necessário na análise matemática subseqüente. Deve ser notado, entretanto, que conforme o diagrama de bloco é simplificado, os novos blocos tornam-se mais complexos devido à geração de novos pólos e zeros.

Na simplificação de um diagrama de blocos, lembre o seguinte:

1. O produto das funções de transferência no sentido direto deve permanecer o mesmo.
2. O produto das funções de transferência em um laço deve permanecer o mesmo.

Uma regra geral para simplificar um diagrama de blocos é mover pontos de

junção e pontos de soma, permutar pontos de soma e, então, reduzir laços de realimentação internos.

Como um exemplo do uso das regras na Tabela 4.3, considere o sistema indicado na Fig. 4.20(a). Movendo o ponto de soma do laço de realimentação negativa contendo H_2 para fora do laço de realimentação positiva contendo H_1, obtemos a Fig. 4.20(b). Eliminando o laço de realimentação positiva, obtemos a Fig. 4.20(c). Então, eliminando o laço contendo H_2/G_1, obtemos a Fig. 4.20(d). Finalmente, eliminando o laço de realimentação, obtemos a Fig. 4.20(e).

Note que o numerador da função de transferência em malha-fechada $C(s)/R(s)$ é o produto das funções de transferência do ramo direto. O denominador de $C(s)/R(s)$ é igual a

$$1 - \sum \text{(produto das funções de transferência em cada laço)}$$

$$= 1 - (G_1G_2H_1 - G_2G_3H_2 - G_1G_2G_3)$$
$$= 1 - G_1G_2H_1 + G_2G_3H_2 - G_1G_2G_3$$

(O laço de realimentação positiva fornece um termo negativo no denominador.)

4.5 DEDUÇÃO DE FUNÇÕES DE TRANSFERÊNCIA DE SISTEMAS FÍSICOS

Sistemas de controle podem consistir em componentes de diferentes tipos tais como elétricos, mecânicos, hidráulicos, pneumáticos ou térmicos. Um engenheiro de controle deve estar familiarizado com as leis fundamentais que envolvem estes componentes.

Nas Seções 4.2 e 4.3, deduzimos funções de transferência para alguns sistemas. Nesta seção apresentaremos exemplos adicionais indicando a dedução das funções de transferência para vários tipos de sistemas físicos.

Na dedução de funções de transferência note o seguinte:

1. Quando aproximar sistemas físicos por modelos lineares a parâmetros concentrados são necessárias certas considerações. Em geral, estas hipóteses dependem das faixas de operação do sistema. Hipóteses que podem ser convenientes para certas condições de operação podem não ser adequadas para outras condições. Na prática, fazer hipóteses apropriadas é uma parte muito importante da análise de um sistema desde que a precisão dos resultados depende da validade dessas considerações. Nesta conexão o engenheiro deve ter em mente que embora as funções de transferência obtidas analiticamente mostrem o efeito geral de variações em parâmetros do sistema na solução da equação, as predições numéricas obtidas a partir de estudos analíticos de sistemas físicos podem não ser precisas. Em outras palavras, embora características qualitativas da dinâmica do sistema possam ser obtidas claramente a partir da função de transferência, a informação quantitativa pode não ser necessariamente precisa. Se for necessária precisão quantitativa, será necessário desenvolver estudos experimentais, além dos analíticos.

2. É desejável verificar a validade da função de transferência admitindo que certos parâmetros do sistema sejam zero ou infinito. Desde que este procedimento simplifica a função de transferência, sua validade pode ser facilmente verificada.

As funções de transferência obtidas a seguir são baseadas na hipótese de que não há efeito de carga na saída. (Lembre-se que se houver qualquer efeito de carga, deve ser levado em conta quando a função de transferência é deduzida.)

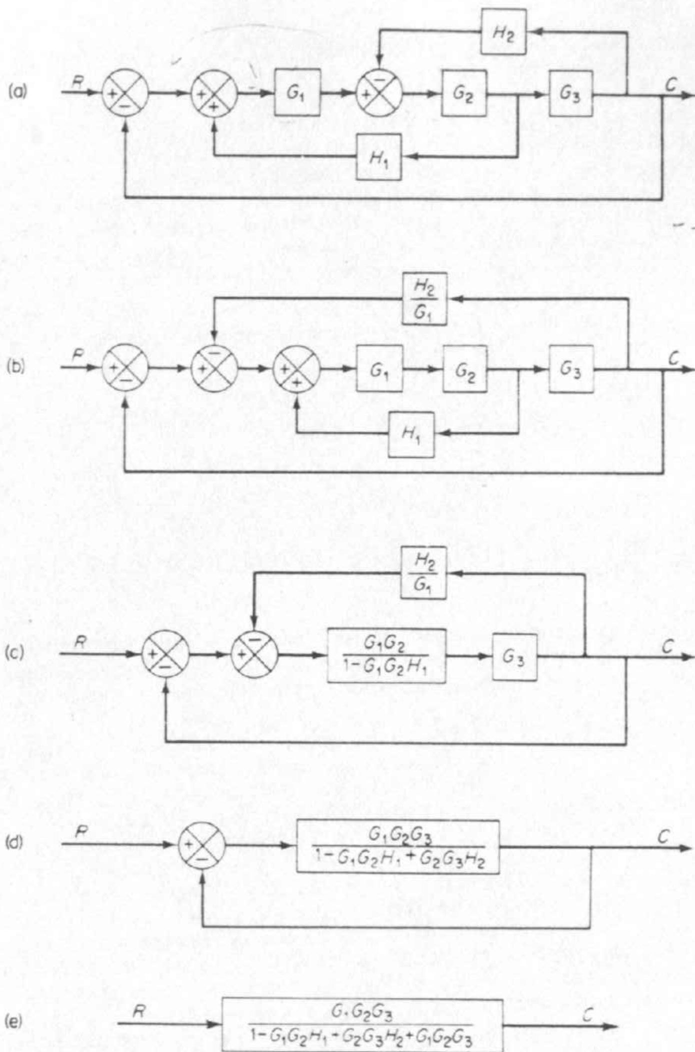


Fig. 4.20 (a) Sistemas de múltiplos laços: (b)-(e) redução sucessiva do diagrama de blocos indicado em (a).

Sistemas Mecânicos e Eletromecânicos

Sismógrafo. A Fig. 4.21 mostra um diagrama esquemático de um sismógrafo. Um sismógrafo indica o deslocamento de sua carcaça em relação ao espaço inercial. É utilizada para medir deslocamentos de terra durante terremotos (abalos sísmicos).

Vamos definir

x_i = deslocamento da carcaça relativo ao espaço inercial
x_0 = deslocamento da massa m relativa ao espaço inercial
$y = x_0 - x_i$ = deslocamento da massa m relativamente à carcaça

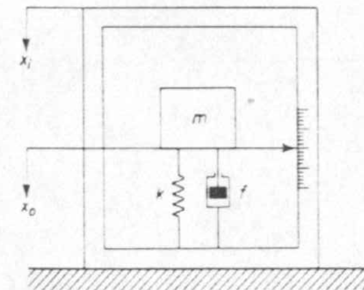


Fig. 4.21 Diagrama esquemático de um sismógrafo.

(Note que, desde que há a produção de uma deflexão estacionária na mola devido à gravidade, medimos o deslocamento x_0 da massa m em relação à posição de equilíbrio estático.) A equação para este sistema é dada por

$$m\ddot{x}_0 + f(\dot{x}_0 - \dot{x}_i) + k(x_0 - x_i) = 0$$

Substituindo $x_0 = y + x_i$ nesta última equação, obtemos uma equação diferencial em y. (Note que y é um sinal que podemos realmente medir.)

$$m\ddot{y} + f\dot{y} + ky = -m\ddot{x}_i \qquad (4.31)$$

Tomando a transformada de Laplace da Eq. (4.31), supondo condições iniciais nulas, obtemos

$$[ms^2 + fs + k]Y(s) = -ms^2X_i(s)$$

Considerando x_i como entrada e y como saída, a função de transferência é

$$\frac{Y(s)}{X_i(s)} = \frac{-ms^2}{ms^2 + fs + k} = -\frac{s^2}{s^2 + \frac{f}{m}s + \frac{k}{m}}$$

Para entradas em baixa potência, a massa m segue a carcaça para cima e para baixo, e a deflexão da mola é muito pequena. Se a entrada x_i consistir em sinais cujas freqüências são muito menores do que $\sqrt{k/m}$, então a função de transferência pode ser aproximada por

$$\frac{Y(j\omega)}{X_i(j\omega)} = \frac{\omega^2}{\frac{k}{m}} \tag{4.32}$$

[Note que na Eq. (4.32), s foi substituído por $j\omega$. Esta função de transferência é denominada função de transferência senoidal. Para detalhes, referir-se ao Cap. 9.]

Como um exemplo, se $y(t)$ é verificado ser senoidal da forma A sen ωt, onde $\omega << \sqrt{k/m}$, então do registro de $y(t)$ o sinal de entrada $x_i(t)$ pode ser determinado como $(A/\omega^2)(k/m)$ sen ωt.

Desde que em sismógrafos a freqüência natural não amortecida $\sqrt{k/m}$ é feita pequena, se a freqüência de entrada é muito maior do que $\sqrt{k/m}$, a massa m permanecerá quase fixa no espaço inercial e um movimento da carcaça indicará o movimento relativo entre a carcaça e a massa. Se o sinal de entrada x_i possuir uma freqüência muito alta, então a função de transferência pode ser aproximada por

$$\frac{Y(j\omega)}{X_i(j\omega)} = -1$$

Isto implica que se $y(t)$ é verificado ser A sen ωt, onde $\omega >> \sqrt{k/m}$, então a entrada $x_i(t)$ pode ser determinada como $-A$ sen ωt.

Servomotores. Os servomotores aqui considerados serão servomotores bifásicos, motores cc controlados por armadura, e motores cc controlados pelo campo. Consideraremos inicialmente o efeito de carga na dinâmica do servomotor.

Efeito de carga na dinâmica de servomotores. A mais importante das características do servomotor é a máxima aceleração obtenível. Para um dado torque disponível o momento de inércia do rotor deve ser um mínimo. Desde que o servomotor opera sob condições continuamente variáveis, ocorrem aceleração e freamento de instante a instante. O servomotor deve ser apto a absorver energia mecânica bem como gerá-la. O desempenho do servomotor, quando utilizado como um freio, deve ser satisfatório.

Seja J_m e f_m, respectivamente, o momento de inércia e a fricção do rotor, e seja J_L e f_L, respectivamente, o momento de inércia e a fricção da carga no eixo de saída. Suponha que o momento de inércia e a fricção do conjunto de engrenagens ou sejam desprezíveis ou estejam incluídos em J_L e f_L, respectivamente. Então, o momento de inércia J_{eq} referido ao eixo do motor e a fricção equivalente f_{eq} referida ao eixo do motor podem ser escritos como (para detalhes vide Problema A.4.2)

$$J_{eq} = J_m + n^2 J_L \qquad (n < 1)$$
$$f_{eq} = f_m + n^2 f_L \qquad (n < 1)$$

onde n é a relação de engrenagens entre o eixo do motor e a carga. Se a relação de engrenagem n é pequena e $J_m >> n^2 J_L$, então o momento de inércia da carga referido ao eixo do motor é desprezível em relação ao momento de inércia do motor. Um argumento similar aplica-se à fricção da carga. Em geral, quando a relação de engrenagem n é pequena, a função de transferência do servomotor elétrico pode ser obtida sem levar-se em consideração o momento de inércia e a fricção da carga. Se nem J_m nem $n^2 J_L$ são desprezíveis, comparados com o outro, entretanto, então o momento de inércia equivalente J_{eq} deve ser utilizado no cálculo da função de transferência do conjunto motor-carga.

Servomotores bifásicos. Um servomotor bifásico, comumente usado em servomecanismos de instrumentação, é análogo a um motor de indução bifásico convencional exceto por suas considerações especiais de projeto. Ele utiliza um rotor em gaiola. Este rotor possui uma pequena relação entre diâmetro-altura a fim de minimizar o momento de inércia e obter uma boa característica de aceleração. O servomotor bifásico é muito forte e confiável.

Em muitas aplicações práticas, a faixa de potência na qual são usados servomotores bifásicos está entre uma fração de um watt e uma centena de watts.

Um diagrama esquemático de um servomotor bifásico é indicado na Fig. 4.22(a). Neste diagrama, uma fase (campo fixo) do motor é excitada continuamente através de uma tensão de referência, cuja freqüência normalmente é 60, 400 ou 1.000 Hz; e a outra fase (campo de controle) é excitada com uma tensão de controle (um sinal com portadora suprimida) que está 90° defasada no tempo em relação à tensão de referência. (A tensão de controle é de amplitude e polaridade variáveis.)

Note que a tensão da fase de controle é feita 90° defasada em relação à tensão da fase fixa. Os enrolamentos do estator para as fases fixa e de controle estão posicionados 90° deslocados no espaço. Estas considerações são baseadas no fato de que o torque é produzido mais eficientemente em um eixo quando os eixos do enrolamento de fase estão em quadratura no espaço e as tensões nas duas fases estão em quadratura no tempo.

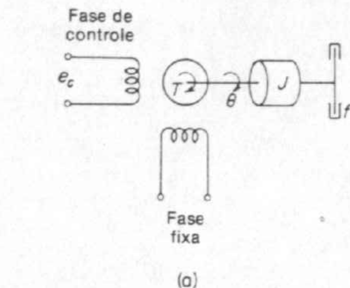
Os dois enrolamentos do estator são normalmente excitados por uma fonte de alimentação bifásica. Se não for disponível uma fonte de alimentação bifásica, entretanto, então o enrolamento de fase fixa pode ser ligado a uma fonte de potência monofásica através de um capacitor, que fornecerá a defasagem de 90°. O amplificador no qual o enrolamento de fase de controle é ligado e alimentado pela mesma fonte de alimentação monofásica.

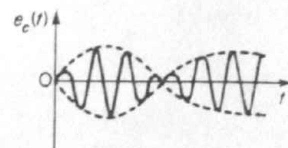
No servomotor bifásico, a polaridade da tensão de controle determina o sentido de rotação. A tensão de controle instantânea $e_c(t)$ é da forma

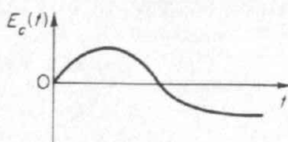
$$\begin{aligned} e_c(t) &= E_c(t) \operatorname{sen} \omega t & \text{para} \quad E_c(t) > 0 \\ &= |E_c(t)| \operatorname{sen}(\omega t + \pi) & \text{para} \quad E_c(t) < 0 \end{aligned}$$

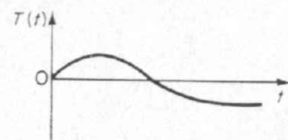
Isto significa que uma variação no sinal de $E_c(t)$ desloca a fase por π radianos. Conseqüentemente, a mudança do sinal da tensão de controle $E_c(t)$ inverte o sentido de rotação do motor. Desde que a tensão de referência é constante, o torque T e a velocidade angular $\dot{\theta}$ também são funções da tensão de controle $E_c(t)$. Se as variações em $E_c(t)$ são lentas comparadas com a freqüência da fonte c.a., o torque desenvolvido pelo motor é proporcional a $E_c(t)$. A Fig. 4.22(b) mostra as curvas $e_c(t)$ *versus* t, $E_c(t)$ *versus* t e o torque $T(t)$ *versus* t. A velocidade angular em regime permanente é proporcional à tensão de controle $E_c(t)$.

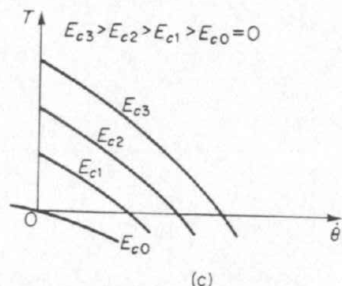
Uma família de curvas torque-velocidade, quando a tensão nominal é aplicada ao enrolamento de fase fixa e várias tensões são aplicadas ao enrolamento de fase de controle, fornece as características em regime estacionário do servomotor bifásico. A função de transferência de um servomotor bifásico pode ser obtida destas curvas torque-velocidade se elas forem retas paralelas e eqüidistantes. Geralmente, as curvas torque-velocidade são paralelas para uma faixa de velocidades relativamente ampla, porém podem não ser eqüidistantes: isto é, para uma dada velocidade, o torque pode não variar linearmente em relação à tensão de controle. Em uma região de baixa velocidade, entretanto, as curvas torque-velocidade são normalmente retas e eqüidistantes em uma região com pequenas tensões de controle. Considerando que o servomotor raramente opera em altas velocidades, as partes lineares das curvas torque-velocidade podem ser estendidas para a região de alta











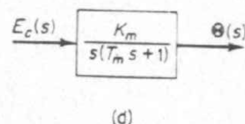


Fig. 4.22 (a) Diagrama esquemático de um servomotor bifásico; (b) curvas mostrando $e_c(t)$ *versus* t, $E_c(t)$ *versus* t e $T(t)$ *versus* t; (c) curvas torque-velocidade; (d) diagrama de blocos de um servomotor bifásico.

velocidade. Se se admite a hipótese de que as curvas são eqüidistantes para todas as tensões de controle, então o servomotor pode ser considerado linear.

A Fig. 4.22(c) mostra um conjunto de curvas torque-velocidade para vários valores de tensões de controle. A curva torque-velocidade correspondente à tensão de controle nula passa pela origem. Desde que a inclinação desta curva normalmente é negativa, se a tensão da fase de controle torna-se igual a zero, o motor desenvolve aquele torque necessário para parar a rotação.

O servomotor fornece um grande torque na velocidade nula. Este torque é necessário para aceleração rápida. Da Fig. 4.22(c), verificamos que o torque T gerado é uma função da velocidade angular $\dot{\theta}$ do eixo do motor e da tensão de controle E_c. A equação para a curva torque-velocidade é

$$T = -K_n\dot{\theta} + K_cE_c \tag{4.33}$$

onde K_n e K_c são constantes positivas. A equação de equilíbrio de torque para o servomotor bifásico é

$$T = J\ddot{\theta} + f\dot{\theta} \tag{4.34}$$

onde J é o momento de inércia do motor e a carga referida ao eixo do motor e f é o coeficiente de fricção-viscosa do motor e da carga referida ao eixo do motor. Das Eqs. (4.33) e (4.34) obtemos a seguinte equação:

$$J\ddot{\theta} + (f + K_n)\dot{\theta} = K_cE_c$$

Notando que a tensão de controle E_c é a entrada e o deslocamento do eixo do motor é a saída, verificamos que a função de transferência do sistema é dada por

$$\frac{\Theta(s)}{E_c(s)} = \frac{K_c}{Js^2 + (f + K_n)s} = \frac{K_m}{s(T_ms + 1)} \tag{4.35}$$

onde

$K_m = K_c/(f + K_n)$ = constante de ganho do motor
$T_m = J/(f + K_n)$ = constante de tempo do motor

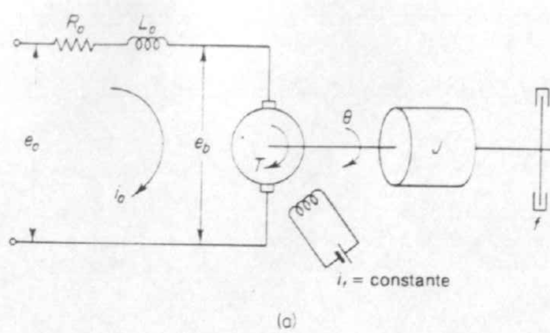
A Fig. 4.22(d) mostra um diagrama de blocos para este sistema. A partir da função de transferência deste sistema, podemos ver que $(f + K_n)s$ é um termo de fricção-viscosa produzido pelo motor e carga. Portanto, K_n, o negativo da inclinação da curva torque-velocidade, junto com f define a fricção-viscosa equivalente do conjunto motor e carga. Para curvas torque-velocidade mais inclinadas, o amortecimento do motor é maior. Se a inércia do rotor é suficientemente baixa, então, em grande parte da faixa de freqüência temos $|T_ms| << 1$ e o servomotor atua como um integrador.

A função de transferência dada pela Eq. (4.35) é baseada na hipótese de que o servomotor é linear. Na prática, entretanto não é exatamente assim. Para curvas torque-velocidade não exatamente paralelas e eqüidistantes, o valor de K_n não é constante e, portanto, os valores de K_m e T_m também não são constantes; eles variam com a tensão de controle.

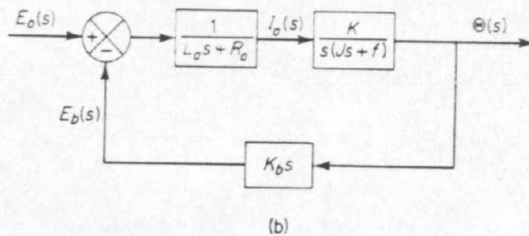
Motores c.c. controlados por armadura. Um motor c.c. é empregado em um sistema de controle quando é exigida uma quantidade apreciável de potência no eixo. Os motores c.c. são muito mais eficientes do que os servomotores c.a. bifásicos.

Os motores c.c. possuem campos excitados separadamente. São ou controlados por armadura com campo fixo ou controlados por campo com corrente de armadura fixa. Por exemplo, motores c.c. usados em instrumentos empregam um campo por magneto-permanente fixo, e o sinal de controle é aplicado aos terminais da armadura.

As características de desempenho do motor c.c. controlado por armadura assemelham-se às características idealizadas do servomotor c.a. bifásico.



(a)



(b)

Fig. 4.23 (a) Diagrama esquemático de um motor c.c. controlado por armadura; (b) diagrama de blocos.

Considere o motor c.c. controlado por armadura indicado na Fig. 4.23(a). Neste sistema,

R_a = resistência do enrolamento da armadura, ohms
L_a = indutância do enrolamento da armadura, henrys
i_a = corrente do enrolamento da armadura, ampères
i_f = corrente de campo, ampères
e_a = tensão aplicada na armadura, volts
e_b = força contra eletromotriz, volts
θ = deslocamento angular do eixo do motor, radianos
T = torque fornecido pelo motor, N · m
J = momento de inércia equivalente do motor e da carga referida ao eixo do motor, kg-m²
f = coeficiente de fricção-viscosa equivalente do motor e da carga referida ao eixo do motor, kg-m/rad/s

O torque T fornecido pelo motor é proporcional ao produto da corrente de armadura i_a e o fluxo no entreferro de ar ψ, que por sua vez é proporcional à corrente de campo ou

$$\psi = K_f i_f$$

onde K_f é uma constante. O torque T pode, portanto, ser escrito

$$T = K_f i_f K_1 i_a$$

onde K_1 é a constante.

No motor c.c. controlado por armadura, a corrente de campo é mantida constante. Para uma corrente de campo constante, resulta um fluxo constante, e o torque torna-se diretamente proporcional à corrente de armadura, de modo que

$$T = K i_a$$

onde K é a constante de torque do motor. Quando a armadura está girando, é induzida na armadura uma tensão proporcional ao produto do fluxo e da velocidade angular. Para um fluxo constante, a tensão induzida e_b é diretamente proporcional à velocidade angular $d\theta/dt$. Portanto,

$$e_b = K_b \frac{d\theta}{dt} \tag{4.36}$$

onde K_b é a constante de força-contra-eletromotriz.

A velocidade de um motor c.c. controlado por armadura é controlada pela tensão de armadura e_a. A tensão de armadura e_a é suprida por um amplificador (ou por um gerador, que é suprido por um amplificador). A equação diferencial para o circuito de armadura é

$$L_a \frac{di_a}{dt} + R_a i_a + e_b = e_a \tag{4.37}$$

A corrente de armadura produz o torque que é aplicado à inércia e à fricção; portanto,

$$J \frac{d^2\theta}{dt^2} + f \frac{d\theta}{dt} = T = K i_a \tag{4.38}$$

Supondo que todas as condições iniciais são nulas e considerando as transformadas de Laplace das Eqs. (4.36), (4.37) e (4.38), obtemos as seguintes equações:

$$K_b s\Theta(s) = E_b(s) \tag{4.39}$$

$$(L_a s + R_a)I_a(s) + E_b(s) = E_a(s) \tag{4.40}$$

$$(Js^2 + fs)\Theta(s) = T(s) = KI_a(s) \tag{4.41}$$

Considerando $E_a(s)$ como a entrada e $\Theta(s)$ como a saída podemos construir os diagramas de blocos das Eqs. (4.39), (4.40) e (4.41), como indicado na Fig. 4.23(b). O efeito da força-contra-eletromotriz é visto como sendo o sinal de realimentação proporcional à velocidade do motor. Esta força-contra-eletromotriz aumenta o amortecimento efetivo do sistema. A função de transferência deste sistema é dada por

$$\frac{\Theta(s)}{E_a(s)} = \frac{K}{s[L_a J s^2 + (L_a f + R_a J)s - R_a f + KK_b]} \tag{4.42}$$

A indutância L_a no circuito de armadura normalmente é pequena e pode ser desprezada. Se L_a for desprezada, então a função de transferência dada pela Eq. (4.42) reduz-se a

$$\frac{\Theta(s)}{E_a(s)} = \frac{K_m}{s(T_m s + 1)} \tag{4.43}$$

onde

$$K_m = K/(R_a f + KK_b) = \text{constante de ganho do motor}$$
$$T_m = R_a J/(R_a f + KK_b) = \text{constante de tempo do motor}$$

A partir das Eqs. (4.42) e (4.43), pode ser visto que as funções de transferência envolvem o termo $1/s$. Portanto, este sistema possui uma propriedade de integração. Na Eq. (4.43), note que a constante de tempo do motor é menor para um R_a menor e J menor. Com J pequeno, conforme a resistência R_a é reduzida, a constante de tempo do motor se aproxima de zero, e o motor atua como um integrador ideal.

Exemplo 4.1 Servomecanismo de posição. Obtenha a função de transferência em malha-fechada para o servomecanismo de posição mostrado na Fig. 4.24. Suponha que a entrada e a saída do sistema sejam a posição do eixo de entrada e a posição do eixo de saída, respectivamente. Suponha os seguintes valores numéricos para as constantes do sistema:

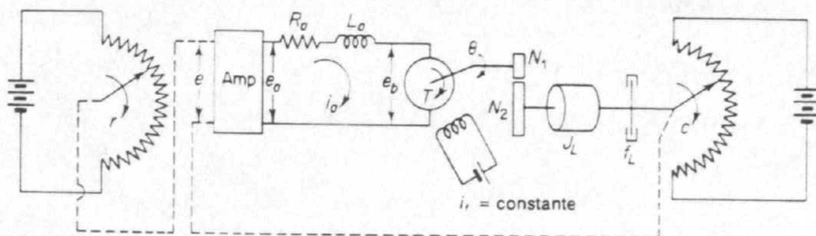


Fig. 4.24 Servomecanismo de posição.

r = deslocamento angular do eixo de entrada de referência, radianos
c = deslocamento angular do eixo de saída, radianos
θ = deslocamento angular do eixo do motor, radianos
K_1 = ganho do detector de erro potenciométrico = $24/\pi$ volts/rad
K_p = ganho do amplificador = 10 volts/volt
e_a = tensão aplicada na armadura, volts
e_b = força-contra-eletromotriz, volts
R_a = resistência do enrolamento da armadura = 0,2 ohms
L_a = indutância do enrolamento da armadura = desprezível
i_a = corrente do enrolamento da armadura, ampères
K_b = força-contra-eletromotriz = $5{,}5 \times 10^{-2}$ volts-s/rad
K = constante de torque do motor = 6×10^{-5} lb-ft/amp
J_m = momento de inércia do motor = 1×10^{-5} lb-ft-s²
f_m = coeficiente de fricção viscosa do motor = desprezível
J_L = momento de inércia da carga = $4{,}4 \times 10^{-3}$ lb-ft-s²
f_L = coeficiente de fricção viscosa da carga = 4×10^{-2} lb-ft/rad/s
n = relação de engrenagem $N_1/N_2 = 1/10$

As equações que descrevem a dinâmica do sistema são as seguintes:
Para o detector de erro potenciométrico:

$$E(s) = K_1[R(s) - C(s)] = 7{,}64[R(s) - C(s)] \tag{4.44}$$

Para o amplificador:

$$E_a(S) = K_p E(s) = 10E(s) \tag{4.45}$$

Para o motor c.c. controlado por armadura: O momento de inércia equivalente J e a fricção viscosa equivalente f referidos ao eixo do motor são, respectivamente,

$$J = J_m + n^2 J_L$$
$$= 1 \times 10^{-5} + 4{,}4 \times 10^{-5} = 5{,}4 \times 10^{-5}$$
$$f = f_m + n^2 f_L$$
$$= 4 \times 10^{-4}$$

Utilizando a Eq. (4.43), obtemos

$$\frac{\Theta(s)}{E_a(s)} = \frac{K_m}{s(T_m s + 1)}$$

onde

$$K_m = \frac{K}{R_a f + KK_b} = \frac{6 \times 10^{-5}}{(0{,}2)(4 \times 10^{-4}) + (6 \times 10^{-5})(5{,}5 \times 10^{-2})} = 0{,}72$$

$$T_m = \frac{R_a J}{R_a f + KK_b} = \frac{(0{,}2)(5{,}4 \times 10^{-5})}{(0{,}2)(4 \times 10^{-4}) + (6 \times 10^{-5})(5{,}5 \times 10^{-2})} = 0{,}13$$

Portanto,

$$\frac{\Theta(s)}{E_a(s)} = \frac{10C(s)}{E_a(s)} = \frac{0{,}72}{s(0{,}13s + 1)} \tag{4.46}$$

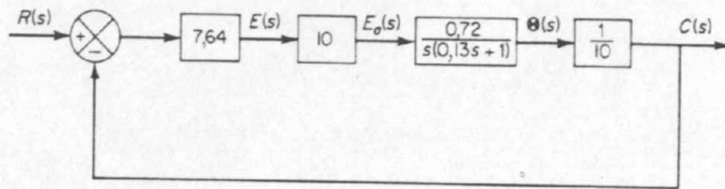
Usando as Eqs. (4.44), (4.45) e (4.46), podemos desenhar o diagrama de blocos do sistema conforme indicado na Fig. 4.25(a). Simplificando este diagrama de blocos, obtemos o indicado na Fig. 4.25(b). A função de transferência em malha-fechada para este sistema é

$$\frac{C(s)}{R(s)} = \frac{42{,}3}{s^2 + 7{,}7s + 42{,}3}$$

Motor c.c. controlado por campo. A Fig. 4.26(a) apresenta um diagrama esquemático de um motor c.c. controlado por campo. Vamos definir

R_f = resistência do enrolamento de campo, ohms
L_f = indutância do enrolamento de campo, henrys
i_f = corrente do enrolamento de campo, ampères
e_f = tensão aplicada de campo, volts
R_a = soma da resistência de armadura e da resistência inserida, ohms
i_a = corrente de armadura, ampères
θ = deslocamento angular do eixo do motor, radianos

T = torque desenvolvido pelo motor, N-m
J = momento de inércia equivalente do motor e da carga referida no eixo do motor, kg-m²
f = coeficiente de fricção-viscosa equivalente do motor e da carga referida ao eixo do motor, kg-m/rad/s



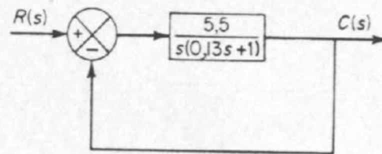
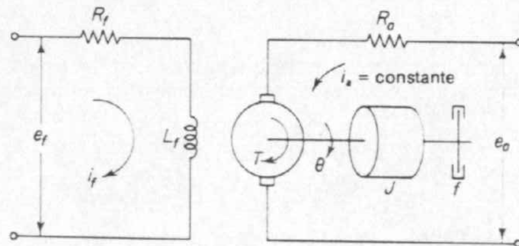


Fig. 4.25 (a) Diagrama de blocos do sistema indicado na Fig. 4.24; (b) diagrama de blocos simplificado.



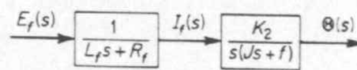


Fig. 4.26 (a) Diagrama esquemático de um motor c.c. controlado por campo; (b) diagrama de blocos.

Neste sistema, a tensão de campo e_f é a entrada de controle. É a saída de um amplificador. A corrente de armadura i_a é mantida constante; esta situação pode ser atingida aplicando-se uma fonte de tensão constante à armadura e uma resistência muito grande em série com a armadura. Se a tensão elétrica nesta resistência for muito maior do que a força-contra-eletromotriz máxima induzida pela rotação dos enrolamentos da armadura no campo magnético, o efeito da força-contra-eletromotriz é desprezível. Deste modo a corrente de armadura i_a pode ser mantida aproximadamente constante. O rendimento do motor neste tipo de operação é necessariamente baixo, porém um motor c.c. controlado pelo campo pode ser utilizado para um sistema de controle de velocidade. Note que é mais difícil manter a corrente de armadura constante do que uma corrente de campo constante, devido à força-contra-eletromotriz no circuito de armadura.

O torque T desenvolvido pelo motor é proporcional ao produto do fluxo no entreferro de ar ψ e a corrente de armadura i_a de modo que

$$T = K_1 \psi i_a \tag{4.47}$$

onde K_1 é uma constante. Desde que o fluxo no entreferro de ar ψ e a corrente de campo i_f são proporcionais, dentro da faixa de operação usual do motor, e i_a é suposta constante, a Eq. (4.47) pode ser escrita como

$$T = K_2 i_f$$

onde K_2 é uma constante. As equações para este sistema são

$$L_f \frac{di_f}{dt} + R_f i_f = e_f \tag{4.48}$$

$$J\frac{d^2\theta}{dt^2} + f\frac{d\theta}{dt} = T = K_2 i_f \tag{4.49}$$

Calculando as transformadas de Laplace nas Eqs. (4.48) e (4.49), supondo que a condição inicial é nula, obtemos as seguintes equações:

$$(L_f s + R_f)I_f(s) = E_f(s) \tag{4.50}$$

$$(Js^2 + fs)\Theta(s) = K_2 I_f(s) \tag{4.51}$$

Considerando $E_f(s)$ como a entrada e $\Theta(s)$ como a saída, podemos construir um diagrama de blocos das Eqs. (4.50) e (4.51), como indicado na Fig. 4.26(b). Deste diagrama de blocos, a função de transferência deste sistema é obtida como sendo

$$\frac{\Theta(s)}{E_f(s)} = \frac{K_2}{s(L_f s + R_f)(Js + f)} = \frac{K_m}{s(T_f s + 1)(T_m s + 1)} \tag{4.52}$$

onde

$K_m = K_2/(R_f f)$ = constante de ganho do motor
$T_f = L_f/R_f$ = constante de tempo do circuito de campo
$T_m = J/f$ = constante de tempo do elemento inércia-fricção

Desde que a indutância de campo L_f não é desprezível, a função de transferência de um motor c.c. controlado por campo é de terceira ordem.

Comparação entre os desempenhos do motor c.c. controlado por armadura e do motor c.c. controlado por campo. Uma vantagem do controle de campo em um motor c.c. é que o amplificador necessário pode ser simplificado devido à pequena potência exigida no controle de campo. A exigência de uma fonte de corrente constante é, entretanto, uma séria desvantagem de operações controladas por campo. Propiciar uma fonte de corrente constante é muito mais difícil do que uma fonte de tensão constante. A operação controlada por campo possui algumas poucas desvantagens a mais em relação à operação controlada por armadura do motor c.c. No motor c.c. controlado por armadura, a força-contra-eletromotriz atua como um amortecimento; no motor c.c. controlado por campo, entretanto, este não é o caso, e o amortecimento necessário deve ser fornecido pelo motor e pela carga. Devido ao baixo rendimento das operações controladas por campo, o calor gerado na armadura pode constituir-se em um problema.

As constantes de tempo do motor c.c. controlado por campo geralmente são maiores do que as constantes de tempo de um motor comparável controlado por armadura. Na comparação das constantes de tempo entre operações controladas por campo e operações controladas por armadura, entretanto, devemos considerar a constante de tempo do amplificador de potência quando considerarmos o estudo das operações controladas por armadura.

Sistemas de Nível de Líquido

Leis de fluxo de fluidos. Na análise de sistemas envolvendo fluxos de fluidos, verificamos ser necessário distinguir os regimes de fluxo em fluxo laminar e fluxo turbulento, de acordo com o valor do número de Reynolds. Se o número de Reynolds for maior do que aproximadamente 3.000 ~ 4.000, então o fluxo é turbulento. O fluxo é laminar se o número de Reynolds for menor do que aproximadamente 2.000. No caso laminar, o fluxo do fluido ocorre em linhas de fluxo sem turbulência. Sistemas envolvendo fluxo turbulento, na maioria das vezes, têm de ser representados por equações diferenciais não lineares, enquanto sistemas envolvendo fluxo laminar podem ser representados por equações diferenciais lineares. (Processos industriais quase sempre envolvem fluxo de líquidos através de tanques e tubulações. O fluxo nestes processos geralmente é turbulento e não laminar.)

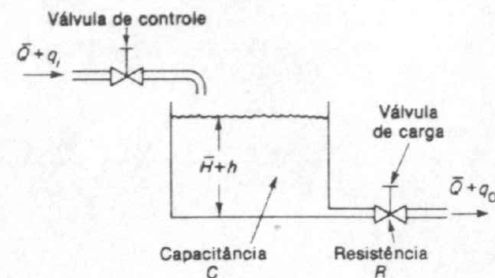
Resistência e capacitância de sistemas de nível de líquido. É conveniente introduzir o conceito de resistência e capacitância para descrever características dinâmicas de sistemas de nível de líquido.

Considere o fluxo através de um pequeno cano interligando dois tanques. A resistência ao fluxo de líquido nesta restrição é definida como a variação na diferença de nível (a diferença entre os níveis de líquido nos dois tanques) necessária para causar uma variação unitária na taxa de fluxo; isto é,

$$R = \frac{\text{variação na diferença de nível, m}}{\text{variação na taxa de fluxo, m}^3/\text{s}}$$

Desde que a relação entre a taxa de fluxo e a diferença de nível difere do fluxo laminar para o fluxo turbulento, consideraremos ambos os casos no que se segue.

Considere o sistema de nível de líquido indicado na Fig. 4.27(a). Neste sistema o líquido flui através da válvula de carga no lado do tanque. Se o fluxo através desta



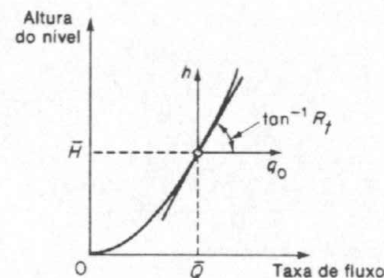


Fig. 4.27 (a) Sistema de nível de líquido; (b) curva da altura de nível *versus* taxa de fluxo.

restrição for laminar, a relação entre a taxa de fluxo em estado estacionário e a altura de líquido no tanque, em estado estacionário no nível de restrição, é dada por

$$Q = KH$$

onde

Q = taxa de fluxo de líquido em estado estacionário, m³/s
K = coeficiente, m²/s
H = altura do nível de líquido em estado estacionário, m

Note que a lei que governa o fluxo laminar é análoga à lei de Coulomb*, que estabelece que a corrente é diretamente proporcional à diferença de potencial.

Para fluxo laminar, a resistência R_ℓ é obtida por

$$R_\ell = \frac{dH}{dQ} = \frac{H}{Q}$$

*N. de T. Na realidade, o autor deve estar referindo-se à lei de Ohm.

A resistência em fluxo laminar é constante e análoga à resistência elétrica.

Se o fluxo através da restrição for turbulento, a taxa de fluxo em estado estacionário é dada por

$$Q = K\sqrt{H} \tag{4.53}$$

onde

Q = taxa de fluxo de líquido em estado estacionário, m^3/s
K = coeficiente, $m^{2,5}/s$
H = altura do nível de líquido em estado estacionário, m

A resistência R_t para fluxo turbulento é obtida de

$$R_t = \frac{dH}{dQ} = \frac{2H}{Q}$$

O valor da resistência em fluxo turbulento depende da taxa de fluxo e da altura do nível.

Pelo uso da resistência em fluxo turbulento, podemos linearizar a relação não linear entre Q e H, conforme indicado na Eq. (4.53). Esta linearização é válida, desde que as variações na altura do nível e a taxa de fluxo, em relação aos seus respectivos valores em estado estacionário, sejam pequenas. A relação linearizada é dada por

$$Q = \frac{2H}{R_t}$$

O valor de R_t pode ser considerado constante se forem pequenas as variações na altura de nível e na taxa de fluxo.

Em muitos casos práticos, o valor do coeficiente K na Eq. (4.53), que depende tanto do coeficiente de fluxo como da área da restrição, não é conhecido. A resistência pode ser então determinada construindo-se o gráfico da curva altura de nível *versus* taxa de fluxo, baseado em dados experimentais, e calculando-se posteriormente a inclinação da curva na condição de operação. Um exemplo de um gráfico deste tipo é indicado na Fig. 4.27(b), e o ponto de operação em regime estacionário e a resistência R_t são indicados na figura. (A resistência R_t é a derivada da curva no ponto de operação.) A aproximação linear é baseada no fato de que a curva real não difere muito da reta tangente se a condição de operação não variar demasiadamente.

A capacitância C de um tanque é definida como sendo a variação na quantidade de líquido armazenado necessária para causar uma variação unitária no potencial (altura do nível de líquido). (O potencial é a grandeza que indica o nível de energia do sistema.)

$$C = \frac{\text{variação no líquido armazenado, m}^3}{\text{variação na altura do nível, m}}$$

Deve ser notado que a capacidade (m^3) e a capacitância (m^2) são diferentes. A capacitância do tanque é igual à sua área da seção transversal. Se esta for constante, a capacitância é constante para qualquer altura do nível de líquido.

Sistemas de nível de líquido. Considere o sistema mostrado na Fig. 4.27(a). As variáveis são definidas como segue:

$\bar{Q}$ = taxa de fluxo em regime estacionário (antes de qualquer variação haver ocorrido), m^3/min
q_i = pequeno desvio da taxa de fluxo de entrada em relação a seu valor em regime estacionário, m^3/min
q_0 = pequeno desvio da taxa de fluxo de saída em relação a seu valor de regime estacionário, m^3/min
$\bar{H}$ = altura do nível em regime estacionário (antes de qualquer variação haver ocorrido), m
h = pequeno desvio na altura do nível em relação a seu valor em regime estacionário, m

Como estabelecido anteriormente, um sistema pode ser considerado linear se o fluxo for laminar. Mesmo que o fluxo seja turbulento, o sistema pode ser linearizado se as variações nas variáveis são mantidas pequenas. Baseado na hipótese de que o sistema ou é linear ou é linearizado, a equação diferencial deste sistema pode ser obtida como segue: desde que o fluxo de entrada menos o fluxo de saída durante o pequeno intervalo de tempo dt é igual à quantidade armazenada adicional no tanque, vemos que

$$C\,dh = (q_i - q_0)\,dt$$

Considerando a definição de resistência, a relação entre q_0 e h é dada por

$$q_0 = \frac{h}{R}$$

A equação diferencial para este sistema para um valor constante de R é a seguinte:

$$RC\frac{dh}{dt} + h = Rq_i \tag{4.54}$$

Note que RC é a constante de tempo do sistema. Considerando as transformadas de Laplace de ambos os lados da Eq. (4.54), supondo uma condição inicial nula, obtemos

$$(RCs + 1)H(s) = RQ_i(s)$$

onde

$$H(s) = \mathscr{L}[h] \quad \text{e} \quad Q_i(s) = \mathscr{L}[q_i]$$

Se q_i é considerada a entrada e h a saída, a função de transferência do sistema é

$$\frac{H(s)}{Q_i(s)} = \frac{R}{RCs + 1}$$

Se, entretanto, q_0 é considerada a saída, com entrada sendo a mesma, então a função de transferência é

$$\frac{Q_0(s)}{Q_i(s)} = \frac{1}{RCs + 1}$$

onde usamos a relação

$$Q_0(s) = \frac{1}{R}H(s)$$

Sistemas de nível de líquido com interação. Considere o sistema mostrado na Fig. 4.28. Neste sistema, os dois tanques interagem. Portanto a função de transferência do sistema não é o produto de duas funções de transferência de primeira ordem.

No que segue, vamos supor apenas pequenas variações das variáveis em relação aos valores de regime estacionário. Usando os símbolos definidos na Fig. 4.28, podemos obter as seguintes equações para este sistema:

$$\frac{h_1 - h_2}{R_1} = q_1 \tag{4.55}$$

$$C_1 \frac{dh_1}{dt} = q - q_1 \tag{4.56}$$

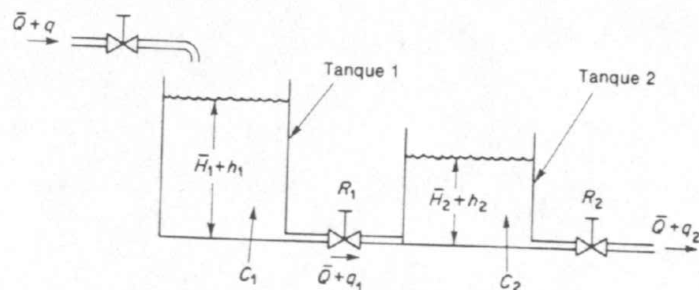
$$\frac{h_2}{R_2} = q_2 \tag{4.57}$$

$$C_2 \frac{dh_2}{dt} = q_1 - q_2 \tag{4.58}$$

Se q é considerada a entrada e q_2 a saída, a função de transferência do sistema é

$$\frac{Q_2(s)}{Q(s)} = \frac{1}{R_1C_1R_2C_2s^2 + (R_1C_1 + R_2C_2 + R_2C_1)s + 1} \tag{4.59}$$

É instrutivo obter a equação (4.59), a função de transferência do sistema interativo, por redução do diagrama de blocos. A partir das Eqs. (4.55) até (4.58), obtemos elementos do diagrama de blocos, conforme mostrado na Fig. 4.29(a). Conectando apropriadamente os sinais, podemos construir um diagrama de blocos conforme indicado na Fig. 4.29(b). Usando as regras da álgebra de diagrama de



$\bar{Q}$: Taxa de fluxo em regime estacionário
$\bar{H}_1$: Nível de líquido do tanque 1 em regime estacionário
$\bar{H}_2$: Nível de líquido do tanque 2 em regime estacionário

Fig. 4.28 Sistema de nível de líquido com interação.

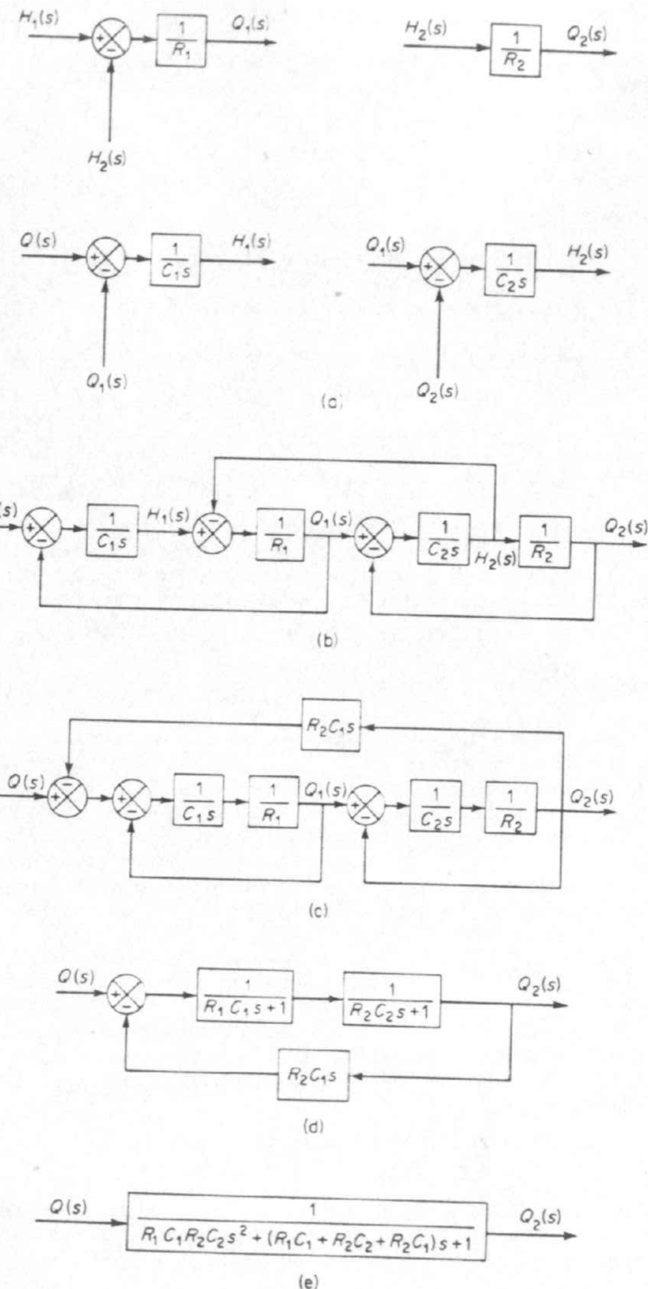


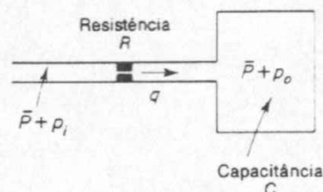
Fig. 4.29 (a) Elementos do diagrama de blocos do sistema indicado na Fig. 4.28; (b) diagrama de blocos do sistema; (c)-(e) redução sucessiva do diagrama de blocos.

blocos fornecidos na Tabela 4.3, este diagrama de blocos pode ser simplificado conforme indicado na Fig. 4.29(c). Simplificações adicionais resultam nas Figs. 4.29(d) e (e). A Fig. 4.29(e) é equivalente à Eq. (4.59).

SISTEMAS DE PRESSÃO

Resistência e capacitância de sistemas de pressão. Muitos processos industriais e controladores pneumáticos envolvem o fluxo de um gás ou ar através de tubulações e recipientes com pressão.

Considere o sistema de pressão mostrado na Fig. 4.30(a). O fluxo de gás através da restrição é uma função da diferença de pressão $p_i - p_0$. Um sistema de



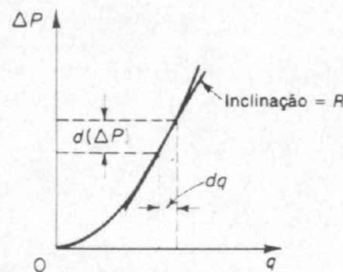


Fig. 4.30 (a) Diagrama esquemático de um sistema de pressão; (b) curva da diferença de pressão *versus* fluxo.

pressão deste tipo pode ser caracterizado em termos de uma resistência e uma capacitância.

A resistência R do fluxo de gás pode ser definida como segue:

$$R = \frac{\text{variação da diferença de pressão de gás, N/m}^2}{\text{variação na taxa de fluxo de gás, kg/s}}$$

ou

$$R = \frac{d(\Delta P)}{dq} \tag{4.60}$$

onde $d\,(\Delta P)$ é uma pequena variação na diferença de pressão do gás e dq uma pequena variação no fluxo de gás. Cálculos do valor da resistência R do fluxo de gás podem ser bastante demorados. Experimentalmente, entretanto, pode ser facilmente determinado através de um gráfico relacionando a diferença de pressão com o fluxo e calculando a inclinação da curva em uma dada condição de operação, conforme mostrado na Fig. 4.30(b).

A capacitância do recipiente sob pressão pode ser definida por

$$C = \frac{\text{variação no gás armazenado, kg}}{\text{variação na pressão do gás, N/m}^2}$$

ou

$$C = \frac{dm}{dp} = V\frac{d\rho}{dp} \tag{4.61}$$

onde

C = capacitância
m = massa de gás no recipiente, kg
p = pressão do gás, N/m²
V = volume do recipiente, m³
ρ = densidade, kg/m³

A capacitância do sistema de pressão depende do tipo de processo de expansão envolvido. A capacitância pode ser calculada pelo uso da lei dos gases ideais. (Vide Problemas A.4.15 até A.4.17.) Se o processo de expansão do gás for politrópico e a variação do estado do gás for entre isotérmica e adiabática, então

$$p\left(\frac{V}{m}\right)^n = \frac{p}{\rho^n} = \text{constante} \tag{4.62}$$

onde

n = expoente politrópico

Para gases ideais,

$$p\bar{v} = \bar{R}T \qquad \text{ou} \qquad pv = \frac{\bar{R}}{M}T$$

onde

p = pressão absoluta, N/m²
$\bar{v}$ = o volume ocupado por 1 mol de um gás, m³/N-mol
$\bar{R}$ = constante universal dos gases, N-m/kg-mol °K
T = temperatura absoluta, °K
v = volume específico do gás, m³/kg
M = massa molecular do gás por mol, kg/kg-mol

Portanto,

$$pv = \frac{p}{\rho} = \frac{\bar{R}}{M}T = R_{gás}T \qquad (4.63)$$

onde

$R_{gás}$ = constante de gás, N-m/kg °K

O expoente politrópico n é unitário para expansão isotérmica. Para expansão adiabática, n é igual à relação entre os calores específicos c_p/c_v, onde c_p é o calor específico a pressão constante e c_v é o calor específico a volume constante. Em muitos casos práticos o valor de n é aproximadamente constante, e portanto a capacitância pode ser considerada constante. O valor de $d\rho/dp$ é obtido das Eqs. (4.62) e (4.63) como

$$\frac{d\rho}{dp} = \frac{1}{nR_{gás}T}$$

A capacitância é então obtida como

$$C = \frac{V}{nR_{gás}T} \qquad (4.64)$$

A capacitância de um dado recipiente é constante se a temperatura permanecer constante. (Em muitos casos práticos, o expoente politrópico n é aproximadamente 1,0 ~ 1,2 para gases em recipientes metálicos não isolados.)

Sistemas de pressão. Considere o sistema mostrado na Fig. 4.30(a). Se supusermos apenas pequenos desvios nas variáveis, em relação a seus valores de regime estacionário, então este sistema pode ser considerado linear.

Vamos definir

$\bar{P}$ = pressão do gás no recipiente em regime estacionário (antes de haverem ocorrido variações na pressão), N/m²
p_i = pequena variação na pressão do gás de entrada, N/m²
p_0 = pequena variação na pressão do gás no recipiente, N/m²
V = volume do recipiente, m³
m = massa do gás no recipiente, kg
q = taxa de fluxo do gás, kg/s
ρ = densidade do gás, kg/s³

Para pequenos valores de p_i e p_0, a resistência R dada pela Eq. (4.60) torna-se constante e pode ser escrita como

$$R = \frac{p_i - p_0}{q}$$

A capacitância C é dada pela Eq. (4.61), reescrita

$$C = \frac{dm}{dp} = V\frac{d\rho}{dp}$$

Desde que a variação de pressão dp_0 vezes a capacitância C é igual ao gás adicionado no recipiente durante dt segundos, obtemos

$$C\,dp_0 = q\,dt$$

ou

$$C\frac{dp_0}{dt} = \frac{p_i - p_0}{R}$$

que pode ser escrita como

$$RC\frac{dp_0}{dt} + p_0 = p_i$$

Se p_i e p_0 são consideradas a entrada e a saída, respectivamente, então a função de transferência do sistema é

$$\frac{P_0(s)}{P_i(s)} = \frac{1}{RCs + 1}$$

onde RC possui a dimensão de tempo e é a constante de tempo do sistema.

SISTEMAS TÉRMICOS

Resistência e capacitância de sistemas térmicos. Sistemas térmicos são aqueles em que há transferência de calor de uma substância para outra. Os sistemas térmicos podem ser analisados em termos de resistência e capacitância, embora a capacitância térmica e a resistência térmica não possam ser representadas precisamente por parâmetros concentrados desde que normalmente são distribuídas através da substância. Para uma análise precisa deveriam ser usados modelos com parâmetros distribuídos. Aqui, entretanto, a fim de simplificar a análise vamos supor que um sistema térmico possa ser representado por um modelo de parâmetros concentrados, no qual as substâncias que são caracterizadas por resistência ao fluxo de calor possuam capacitância ao calor desprezível e as substâncias que são caracterizadas por capacitância térmica (ao calor) possuam resistência térmica desprezível.

Há três diferentes modos pelos quais o calor pode fluir de uma substância para a outra: condução, convecção e radiação.

Para transferência de calor por condução ou convecção,

$$q = K\,\Delta\theta$$

onde

q = taxa de fluxo de calor, cal/s
$\Delta\theta$ = diferença de temperatura, °C
K = coeficiente, cal/s °C

O coeficiente K é dado por

$$K = \frac{kA}{\Delta X} \quad \text{para condução}$$

$$= HA \quad \text{para convecção}$$

onde

k = condutividade térmica, cal/m s °C
A = área normal ao fluxo de calor, m²
ΔX = espessura do condutor, m
H = coeficiente de convecção, cal/m² s °C

Para transferência de calor por radiação, o fluxo de calor é dado por

$$q = K_r(\theta_1^4 - \theta_2^4)$$

onde

q = taxa de fluxo de calor, cal/s
K_r = coeficiente que depende da emissividade, dimensão e configuração da superfície que emana o calor e da superfície que o recebe
θ_1 = temperatura absoluta do emissor, °K
θ_2 = temperatura absoluta do receptor, °K

Desde que a constante K_r é um número muito pequeno, a transferência de calor por radiação somente é apreciável se a temperatura do emissor for muito alta.

A resistência térmica R para a transferência de calor entre duas substâncias pode ser definida como segue:

$$R = \frac{\text{variação na diferença de temperatura, °C}}{\text{variação na taxa de fluxo de calor, cal/s}}$$

A resistência térmica para transferência de calor por condução ou convecção é dada por

$$R = \frac{d(\Delta\theta)}{dq} = \frac{1}{K}$$

Desde que a condutividade térmica e o coeficiente de convecção são praticamente constantes, a resistência térmica tanto para condução como para convecção é constante. A resistência térmica para transferência de calor por radiação é dada por

$$R = \frac{d(\Delta\theta)}{dq} = \frac{1}{4K_r\bar{\theta}^3}$$

onde

$\bar{\theta}$ = diferença de temperatura efetiva do emissor e receptor

A resistência de radiação não é constante; ela varia. Pode ser considerada constante apenas para um pequeno intervalo de condições de operação.

A capacitância térmica C é definida por

$$C = \frac{\text{variação no calor armazenado, cal}}{\text{variação na temperatura, °C}}$$

ou

$$C = Wc_p$$

onde

W = massa da substância considerada, kg
c_p = calor específico da substância, cal/°C-kg

Sistemas térmicos. Considere o sistema mostrado na Fig. 4.31. É suposto que o tanque é isolado para eliminar perda de calor para o ar na vizinhança. Também é suposto que não há armazenamento de calor na isolação e que o líquido no tanque é perfeitamente misturado de modo a estar em uma temperatura uniforme. Conseqüentemente uma temperatura única é usada para descrever a temperatura do líquido no tanque e do líquido saindo do tanque.

Vamos definir

$\bar{\Theta}_i$ = temperatura em regime estacionário do líquido entrando, °C
$\bar{\Theta}_0$ = temperatura em regime estacionário do líquido saindo, °C
G = taxa de fluxo do líquido em regime estacionário, kg/s
M = massa do líquido no tanque, kg
c = calor específico do líquido, cal/kg °C
R = resistência térmica, °Cs/cal
C = capacitância térmica, cal/°C
$\bar{H}$ = taxa de entrada de calor em regime estacionário, cal/°C

Suponha que a temperatura do líquido entrante é mantida constante e que a taxa de entrada de calor é subitamente variada de $\bar{H}$ para $\bar{H} + h_i$, onde h_i representa uma pequena variação na taxa de entrada de calor. A taxa de saída de calor variará então gradualmente desde $\bar{H}$ até $\bar{H} + h_0$. A temperatura do líquido que sai também variará desde $\bar{\Theta}_0$ até $\bar{\Theta}_0 + \theta$. Para este caso, h_0, C e R são obtidas, respectivamente, por

$$h_0 = Gc\theta$$

$$C = Mc$$

$$R = \frac{\theta}{h_0} = \frac{1}{Gc}$$

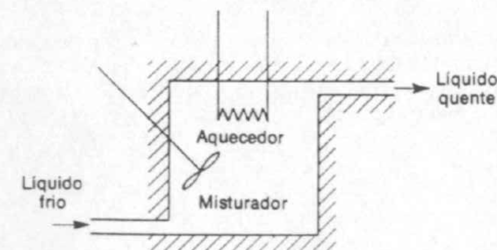


Fig. 4.31 Sistema térmico.

A equação diferencial para este sistema é

$$C\frac{d\theta}{dt} = h_i - h_0$$

que pode ser reescrita como

$$RC\frac{d\theta}{dt} + \theta = Rh_i$$

Note que a constante de tempo do sistema é igual a RC ou M/G segundos. A função de transferência relacionando θ e h_i é dada por

$$\frac{\Theta(s)}{H_i(s)} = \frac{R}{RCs + 1}$$

onde

$$\Theta(s) = \mathcal{L}[\theta(t)] \qquad \text{e} \qquad H_i(s) = \mathcal{L}[h_i(t)]$$

Na prática, a temperatura do líquido que entra pode flutuar e atuar como um distúrbio de carga. (Se uma temperatura de saída constante é desejada, um controlador automático pode ser instalado para ajustar a taxa de fluxo de entrada de calor de maneira a compensar as flutuações na temperatura do líquido de entrada.) Se a temperatura do líquido fluindo na entrada for subitamente variada desde $\bar{\Theta}_i$ até $\bar{\Theta}_i + \theta_i$, enquanto a taxa de entrada de calor H e a taxa de fluxo de líquido G são mantidas constantes, então a taxa de fluxo de saída de calor variará desde $\bar{H}$ para $\bar{H} + h_0$, e a temperatura do líquido que flui na saída variará desde $\bar{\Theta}_0$ para $\bar{\Theta}_0 + \theta$. A equação diferencial para este caso é

$$C\frac{d\theta}{dt} = Gc\theta_i - h_0$$

que pode ser reescrita

$$RC\frac{d\theta}{dt} + \theta = \theta_i$$

A função de transferência relacionando θ e θ_i é dada por

$$\frac{\Theta(s)}{\Theta_i(s)} = \frac{1}{RCs + 1}$$

onde

$$\Theta(s) = \mathcal{L}[\theta(t)] \qquad \text{e} \qquad \Theta_i(s) = \mathcal{L}[\theta_i(t)]$$

Se o sistema térmico apresentado está sujeito tanto a variações na temperatura do líquido que flui na entrada como na taxa de entrada de calor, enquanto a taxa do fluxo de líquido é mantida constante, a variação θ na temperatura do líquido que flui na saída pode ser dada pela seguinte equação:

$$RC\frac{d\theta}{dt} + \theta = \theta_i + Rh_i$$

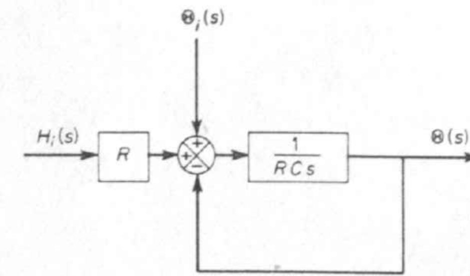


Fig. 4.32 Diagrama de blocos do sistema indicado na Fig. 4.31.

Um diagrama de blocos correspondente a este caso é mostrado na Fig. 4.32. (Note que o sistema envolve duas entradas. Discutiremos sistemas de múltiplas-entradas-múltiplas-saídas na Seção 4.6.)

4.6 SISTEMAS MULTIVARIÁVEIS E MATRIZES DE TRANSFERÊNCIA

Na Seção 4.2, a função de transferência foi definida para sistema de entrada única-saída única. Nesta seção estenderemos a representação de função de transferência a sistemas com múltiplas entradas e múltiplas saídas.

Matrizes de transferência. Considere um sistema com m entradas e n saídas. Podemos considerar as m entradas como sendo os componentes de um vetor. Denominaremos este vetor por vetor de entrada. Analogamente podemos considerar as n saídas como componentes de um vetor de saída. A matriz que relaciona a transformada de Laplace do vetor de saída com a transformada de Laplace do vetor de entrada é denominada a matriz de transferência entre o vetor de saída e o vetor de entrada.

Considere o sistema mostrado na Fig. 4.33. Este sistema tem duas entradas e duas saídas. Da Fig. 4.33, a relação entre as saídas e as entradas é dada por

$$X_1(s) = G_{11}(s)U_1(s) + G_{12}(s)U_2(s)$$
$$X_2(s) = G_{21}(s)U_1(s) + G_{22}(s)U_2(s)$$

onde $G_{ij}(s)$ é a função de transferência relacionando a i-ésima saída com a j-ésima entrada. Utilizando a notação matriz-vetor, podemos escrever esta relação de transferência como sendo

$$\begin{bmatrix} X_1(s) \\ X_2(s) \end{bmatrix} = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \begin{bmatrix} U_1(s) \\ U_2(s) \end{bmatrix}$$

Um sistema possuindo múltiplas entradas e múltiplas saídas é chamado sistema *multivariável*. Se um sistema deste tipo possuir m entradas e n saídas, e se a função de transferência entre a i-ésima saída e a j-ésima entrada for dada por $G_{ij}(s)$,

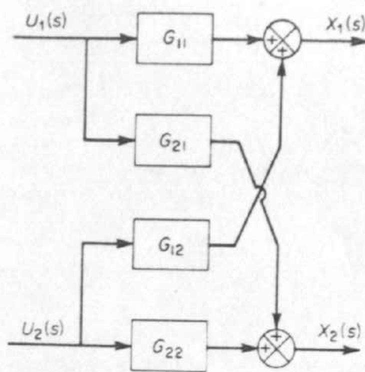


Fig. 4.33 Sistema de múltiplas entradas-múltiplas saídas.

então a transformada de Laplace da i-ésima saída é relacionada com as transformadas de Laplace das m entradas por

$$X_i(s) = G_{i1}(s)U_1(s) + G_{i2}(s)U_2(s) + \cdots + G_{im}(s)U_m(s)$$
$$(i = 1, 2, \ldots, n)$$

Note que na definição de $G_{ij}(s)$, somente a j-ésima entrada é considerada e todas as outras entradas são supostas nulas. Na forma matricial, a transformada de Laplace do vetor de saída é relacionada com a transformada de Laplace do vetor de entrada pela seguinte equação:

$$\begin{bmatrix} X_1(s) \\ X_2(s) \\ \cdot \\ \cdot \\ \cdot \\ X_n(s) \end{bmatrix} = \begin{bmatrix} G_{11}(s) & G_{12}(s) & \cdots & G_{1m}(s) \\ G_{21}(s) & G_{22}(s) & \cdots & G_{2m}(s) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ G_{n1}(s) & G_{n2}(s) & \cdots & G_{nm}(s) \end{bmatrix} \begin{bmatrix} U_1(s) \\ U_2(s) \\ \cdot \\ \cdot \\ \cdot \\ U_m(s) \end{bmatrix} \tag{4.65}$$

A Eq. (4.65) indica as interações entre as m entradas e as n saídas. A Eq. (4.65) pode ser reescrita como

$$\mathbf{X}(s) = \mathbf{G}(s)\mathbf{U}(s)$$

onde

$$\mathbf{X}(s) = \begin{bmatrix} X_1(s) \\ X_2(s) \\ \cdot \\ \cdot \\ \cdot \\ X_n(s) \end{bmatrix}, \qquad \mathbf{U}(s) = \begin{bmatrix} U_1(s) \\ U_2(s) \\ \cdot \\ \cdot \\ \cdot \\ U_m(s) \end{bmatrix}$$

$$\mathbf{G}(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) & \cdots & G_{1m}(s) \\ G_{21}(s) & G_{22}(s) & \cdots & G_{2m}(s) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ G_{n1}(s) & G_{n2}(s) & \cdots & G_{nm}(s) \end{bmatrix}$$

$\mathbf{X}(s)$ é a transformada de Laplace do vetor de saída, $\mathbf{U}(s)$ é a transformada de Laplace do vetor de entrada, e $\mathbf{G}(s)$ é a matriz de transferência entre $\mathbf{X}(s)$ e $\mathbf{U}(s)$.

Exemplo 4.2 Considere o sistema mecânico mostrado na Fig. 4.34. Suporemos que o sistema está inicialmente no repouso.

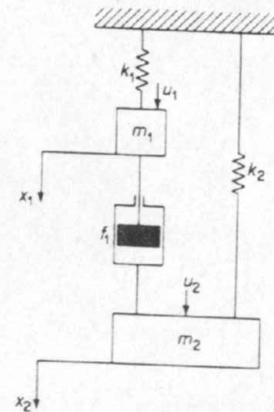


Fig. 4.34 Sistema mecânico.

Este sistema possui duas entradas $u_1(t)$ e $u_2(t)$ e duas saídas $x_1(t)$ e $x_2(t)$. As equações que descrevem a dinâmica do sistema são

$$m_1\ddot{x}_1 + f_1(\dot{x}_1 - \dot{x}_2) + k_1x_1 = u_1$$
$$m_2\ddot{x}_2 + f_1(\dot{x}_2 - \dot{x}_1) + k_2x_2 = u_2$$

Considerando as transformadas de Laplace destas duas equações e substituindo as condições iniciais nulas, obtemos

$$(m_1s^2 + f_1s + k_1)X_1(s) - f_1sX_2(s) = U_1(s)$$
$$(m_2s^2 + f_1s + k_2)X_2(s) - f_1sX_1(s) = U_2(s)$$

Na forma matriz-vetor, obtemos

$$\begin{bmatrix} m_1s^2 + f_1s + k_1 & -f_1s \\ -f_1s & m_2s^2 + f_1s + k_2 \end{bmatrix} \begin{bmatrix} X_1(s) \\ X_2(s) \end{bmatrix} = \begin{bmatrix} U_1(s) \\ U_2(s) \end{bmatrix} \tag{4.66}$$

Vamos definir

$$\Delta = (m_1s^2 + f_1s + k_1)(m_2s^2 + f_1s + k_2) - f_1^2s^2 \neq 0$$

Então, premultiplicando pela inversa da matriz 2 × 2 na Eq. (4.66), obtemos

$$\begin{bmatrix} X_1(s) \\ X_2(s) \end{bmatrix} = \begin{bmatrix} \dfrac{m_2s^2 + f_1s + k_2}{\Delta} & \dfrac{f_1s}{\Delta} \\ \dfrac{f_1s}{\Delta} & \dfrac{m_1s^2 + f_1s + k_1}{\Delta} \end{bmatrix} \begin{bmatrix} U_1(s) \\ U_2(s) \end{bmatrix}$$

A matriz 2 × 2 desta última equação é a matriz de transferência entre as saídas e as entradas. Obviamente, as respostas temporais $x_1(t)$ e $x_2(t)$ são dadas por

$$x_1(t) = \mathscr{L}^{-1}\left[\frac{m_2s^2 + f_1s + k_2}{\Delta} U_1(s) + \frac{f_1s}{\Delta} U_2(s)\right] \tag{4.67}$$

$$x_2(t) = \mathscr{L}^{-1}\left[\frac{f_1s}{\Delta} U_1(s) + \frac{m_1s^2 + f_1s + k_1}{\Delta} U_2(s)\right] \tag{4.68}$$

Para determinar as respostas $x_1(t)$ e $x_2(t)$ para $u_1 = 0$, $u_2 = 0$ (ou $u_1 = 0$, $u_2 \neq 0$), simplesmente substituímos $U_2(s) = 0$ (ou $U_1(s) = 0$) nas Eqs. (4.67) e (4.68).

Exemplo 4.3 O sistema mostrado na Fig. 4.35 possui duas entradas, a entrada de referência e a entrada de distúrbio e uma saída. Obtenha a matriz de transferência entre a saída e as entradas.

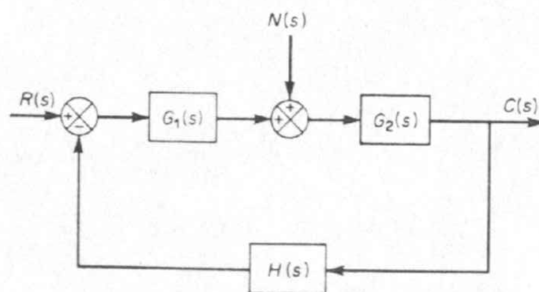


Fig. 4.35 Sistema possuindo duas entradas e uma saída.

A transformada de Laplace da saída $C(s)$ pode ser obtida como

$$C(s) = \frac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)H(s)} R(s) + \frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)} N(s)$$

Na forma matriz-vetor,

$$C(s) = \begin{bmatrix} \dfrac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)H(s)} & \dfrac{G_2(s)}{1 + G_1(s)G_2(s)H(s)} \end{bmatrix} \begin{bmatrix} R(s) \\ N(s) \end{bmatrix} \tag{4.69}$$

A matriz de transferência entre a saída $C(s)$ e as entradas $R(s)$ e $N(s)$ é a matriz 1 × 2 dada na Eq. (4.69).

Comentários. Nesta seção, apresentamos uma definição da matriz de transferência e deduzimos as matrizes de transferência para dois sistemas. A representação por matriz de transferência para sistemas multivariáveis é uma extensão da representação por função de transferência de sistemas de entrada-única-saída-única. A análise e controle ótimo de sistemas multivariáveis pode ser conduzida mais convenientemente pelo uso de variáveis de estado. Conseqüentemente, adiaremos a análise de sistemas multivariáveis para os Caps. 14-16, onde estudaremos a abordagem no espaço de estados para a análise e otimização de sistemas de controle.

4.7 GRÁFICOS DE FLUXO DE SINAL

O diagrama de blocos é útil para representar graficamente sistemas de controle. Entretanto, para um sistema muito complicado, o processo de redução de diagramas de blocos se torna muito demorado. Uma abordagem alternativa para determinar as relações entre as variáveis do sistema, de um sistema de controle complicado, é a abordagem por gráfico de fluxo de sinal, devida a S. J. Mason.

Gráficos de fluxo de sinal. Um gráfico de fluxo de sinal é um diagrama que representa um conjunto de equações algébricas lineares simultâneas. Quando se aplica o método do gráfico de fluxo do sinal para análise de sistemas de controle, devem-se inicialmente transformar as equações diferenciais lineares em equações algébricas em s.

Um gráfico de fluxo de sinal consiste em uma rede na qual os nós são ligados por ramos diretos. Cada nó representa uma variável do sistema e cada ramo ligado entre dois nós atua como um multiplicador de sinal. Note que o sinal flui em apenas um sentido. O sentido do fluxo do sinal é indicado por uma seta colocada no ramo, e o fator de multiplicação é indicado ao longo do ramo. O gráfico de fluxo do sinal indica o fluxo de sinais de um ponto de um sistema para um outro e fornece as relações entre os sinais.

Como é de esperar, um gráfico de fluxo de sinal contém essencialmente a mesma informação do diagrama de blocos. A vantagem do uso de um gráfico de fluxo de sinal para representar um sistema de controle é que uma fórmula de ganho, denominada fórmula de ganho de Mason, é disponível fornecendo as relações entre as variáveis do sistema sem exigir uma redução do gráfico.

Definições. Antes de discutirmos sobre os aspectos de gráficos de fluxo de sinal, devemos definir alguns termos.

Nó. Um nó é um ponto representando uma variável ou um sinal.

Transmitância. A transmitância é um ganho entre dois nós.

Ramo. Um ramo é um segmento de reta orientado unindo dois nós. O ganho de um ramo é uma transmitância.

Nó de entrada ou fonte. Um nó de entrada ou fonte é um nó que possui apenas ramos eferentes (que saem do nó). Corresponde a uma variável independente.

Nó de saída ou sorvedouro. Um nó de saída ou sorvedouro é um nó no qual há apenas ramos aferentes (que entram no nó). Corresponde a uma variável dependente.

Nó misturado (ou misto). Um nó misturado é um nó que possui tanto ramos aferentes como eferentes.

Caminho. Um caminho é uma trajetória de ramos ligados no sentido das flechas dos ramos. Se nenhum nó é cruzado mais de uma vez, o caminho é aberto. Se o caminho termina no mesmo nó no qual começou e não cruza qualquer outro nó mais do que uma vez, ele é fechado. Se um caminho cruza algum nó mais de uma vez porém termina em um nó diferente do qual começou, ele não é nem aberto nem fechado.

Laço. Um laço é um caminho fechado.

Ganho de laço. O ganho de laço é o produto das transmitâncias dos ramos de um laço.

Laços que não se tocam. Laços não se tocam se não possuírem qualquer nó comum.

Caminho direto. Um caminho direto é um caminho desde um nó de entrada (fonte) até um nó de saída (sorvedouro) que não cruza qualquer nó mais do que uma vez.

Ganho do caminho direto. Um ganho do caminho direto é o produto das transmitâncias dos ramos de um caminho direto.

A Fig. 4.36 mostra nós e ramos, além de transmitâncias.

Propriedades dos gráficos de fluxo de sinal. Algumas poucas propriedades importantes de gráficos de fluxo de sinal são alistadas abaixo.

1. Um ramo indica a dependência funcional de um sinal em relação a outro. Um sinal o atravessa apenas no sentido especificado pela seta do ramo.
2. Um nó soma os sinais de todos os ramos aferentes e transmite esta soma a todos os ramos eferentes.
3. Um nó misturado, que possui tanto ramos aferentes como eferentes, pode ser considerado como um nó de saída (sorvedouro) adicionando-se um ramo eferente de transmitância unitária. (Vide Fig. 4.36. Note que o ramo com transmitância unitária é dirigido de x_3 para um outro nó, também denotado por x_3.) Note, entretanto, que não podemos mudar um nó misturado para uma fonte utilizando este método.
4. Para um sistema dado, um gráfico de fluxo de sinal não é único. Muitos gráficos de fluxo de sinal diferentes podem ser desenhados para um dado sistema, escrevendo-se diferentemente as equações do sistema.

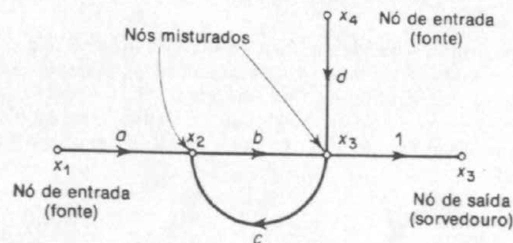


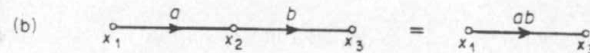
Fig. 4.36 Gráfico de fluxo de sinal.

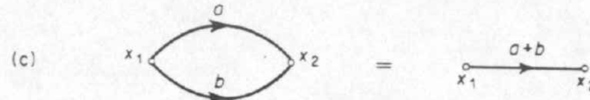
Álgebra do gráfico de fluxo de sinal. Um gráfico de fluxo de sinal de um sistema linear pode ser desenhado usando-se as definições anteriores. Assim fazendo, normalmente colocamos os nós de entrada (fontes) à esquerda e os nós de saída (sorvedouro) à direita. As variáveis independentes e dependentes das equações tornam-se os nós de entrada (fontes) e os nós de saída (sorvedouros), respectivamente. As transmitâncias dos ramos podem ser obtidas a partir dos coeficientes das equações.

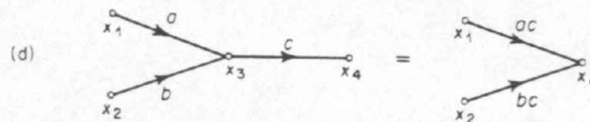
Para determinar a relação entrada-saída, devemos usar a fórmula de Mason, que será fornecida posteriormente, ou podemos reduzir o gráfico de fluxo de sinal a um gráfico contendo apenas os nós de entrada e saída. Para obter este resultado, utilizamos as seguintes regras:

1. O valor de um nó com um ramo aferente, como o indicado na Fig. 4.37(a), é $x_2 = ax_1$.









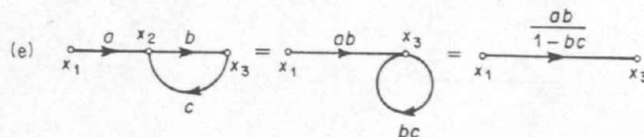


Fig. 4.37 Gráficos de fluxo de sinal e simplificações.

2. A transmitância total de ramos em cascata é igual ao produto das transmitâncias de todos os ramos. Ramos em cascata podem, então, ser combinados em um único ramo pela multiplicação das transmitâncias, conforme mostrado na Fig. 4.37(b).
3. Ramos em paralelo podem ser combinados adicionando-se as transmitâncias conforme a Fig. 4.37(c).
4. Um nó misturado pode ser eliminado conforme a Fig. 4.37(d).
5. Um laço pode ser eliminado conforme indicado na Fig. 4.37(e). Note que

$$x_3 = bx_2, \qquad x_2 = ax_1 + cx_3$$

Portanto,

$$x_3 = abx_1 + bcx_3 \tag{4.70}$$

ou

$$x_3 = \frac{ab}{1 - bc} x_1 \tag{4.71}$$

A Eq. (4.70) corresponde a um diagrama possuindo um autolaço de transmitância bc. A eliminação do autolaço resulta na Eq. (4.71), que mostra claramente que a transmitância global é $ab/(1 - bc)$.

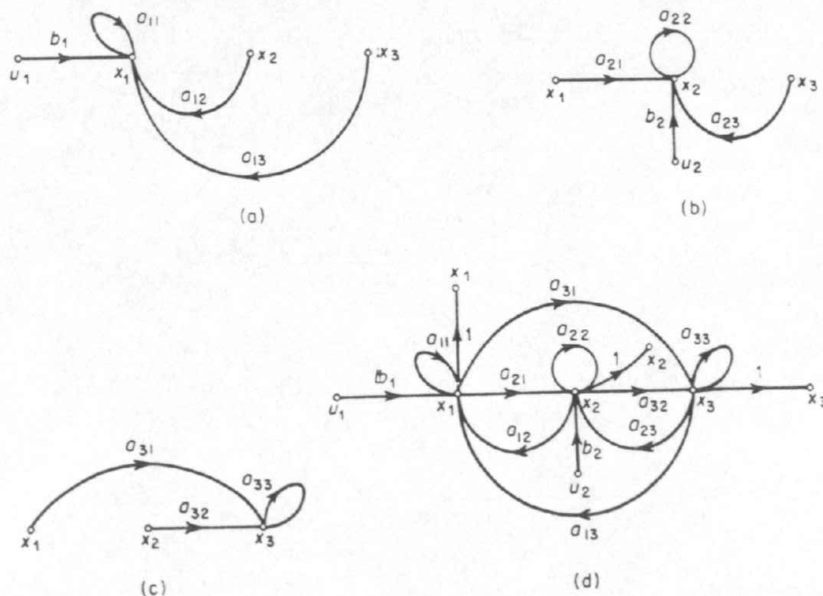


Fig. 4.38 (a) Gráfico de fluxo de sinal representando a Eq. (4.72); (b) gráfico do fluxo de sinal representando a Eq. (4.73); (c) gráfico do fluxo de sinal representando a Eq. (4.74); (d) gráfico de fluxo de sinal completo para um sistema descrito pelas Eqs. (4.72), (4.73), (4.74).

Representação de gráfico de fluxo de sinal para sistemas lineares. Os gráficos de fluxo de sinal são amplamente aplicados na análise de sistemas lineares. Neste caso o gráfico pode ser desenhado a partir das equações do sistema ou, com prática, por inspeção do sistema físico. Redução de rotina pelo uso das regras anteriores fornece a relação entre as variáveis de entrada e saída.

Considere um sistema definido pelo seguinte conjunto de equações:

$$x_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + b_1u_1 \tag{4.72}$$

$$x_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + b_2u_2 \tag{4.73}$$

$$x_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \tag{4.74}$$

onde u_1 e u_2 são variáveis de entrada; x_1, x_2 e x_3 são variáveis de saída. Um gráfico de fluxo de sinal para este sistema, uma representação gráfica dessas três equações simultâneas, indicando a interdependência das variáveis, pode ser obtido como segue: Localize inicialmente os nós x_1, x_2 e x_3, como indicado na Fig. 4.38(a). Note que a_{ij} é a transmitância entre x_j e x_i. A equação (4.72) estabelece que x_1 é igual à soma dos quatros sinais $a_{11}x_1$, $a_{12}x_2$, $a_{13}x_3$ e b_1u_1. O gráfico de fluxo de sinal representando a Eq. (4.72) é indicado na Fig. 4.38(a). A Eq. (4.73) estabelece que x_2 é igual à soma de $a_{21}x_1$, $a_{22}x_2$, $a_{23}x_3$ e b_2u_2. O gráfico de fluxo de sinal correspondente é mostrado na Fig. 4.38(b). O gráfico do fluxo de sinal representando a Eq. (4.74) é indicado na Fig. 4.38(c).

O gráfico de fluxo de sinal representando as Eqs. (4.72), (4.73) e (4.74) é então obtido pela combinação das Figs. 4.38(a), (b) e (c). Finalmente, o gráfico do fluxo de sinal completo para as equações simultâneas fornecidas é mostrado na Fig. 4.38(d).

Em um gráfico de fluxo de sinal, os nós de entrada (fontes) devem ser considerados um de cada vez. O sinal de saída é então igual à soma das contribuições individuais de cada entrada.

O ganho global de uma entrada para uma saída pode ser obtido diretamente do gráfico do fluxo de sinal por inspeção, pelo uso da fórmula de Mason, ou por uma redução do gráfico a uma forma mais simples.

Gráficos de fluxo de sinal para sistemas de controle. Alguns gráficos de fluxo de sinal de sistemas de controle simples são fornecidos na Fig. 4.39. Para estes gráficos simples, a função de transferência em malha-fechada $C(s)/R(s)$ [ou $C(s)/N(s)$] pode ser obtida facilmente por inspeção. Para gráficos de fluxo de sinal mais complicado, a fórmula de ganho de Mason é muito útil.

Fórmula de ganho de Mason para gráficos de fluxo de sinal. Em muitos casos práticos, desejamos determinar a relação entre uma variável de entrada e uma variável de saída do gráfico de fluxo de sinal. A transmitância entre um nó de entrada e um nó de saída é o ganho global, ou transmitância global entre estes dois nós. A fórmula de ganho de Mason, que é aplicável ao ganho global, é dada por

$$P = \frac{1}{\Delta} \sum_k P_k \Delta_k$$

onde

P_k = ganho do caminho ou transmitância do k-ésimo caminho direto.
Δ = determinante do gráfico

= 1 − (soma de todos os ganhos de laço diferentes) + (soma dos produtos de ganhos de todas as possíveis combinações de dois laços que não se tocam) − (soma dos produtos de ganhos de todas as combinações possíveis de três laços que não se tocam) + . . .

$$= 1 - \sum_a L_a + \sum_{b,c} L_b L_c - \sum_{d,e,f} L_d L_e L_f + \cdots$$

$\sum_a L_a$ = soma de todas os ganhos de laços diferentes

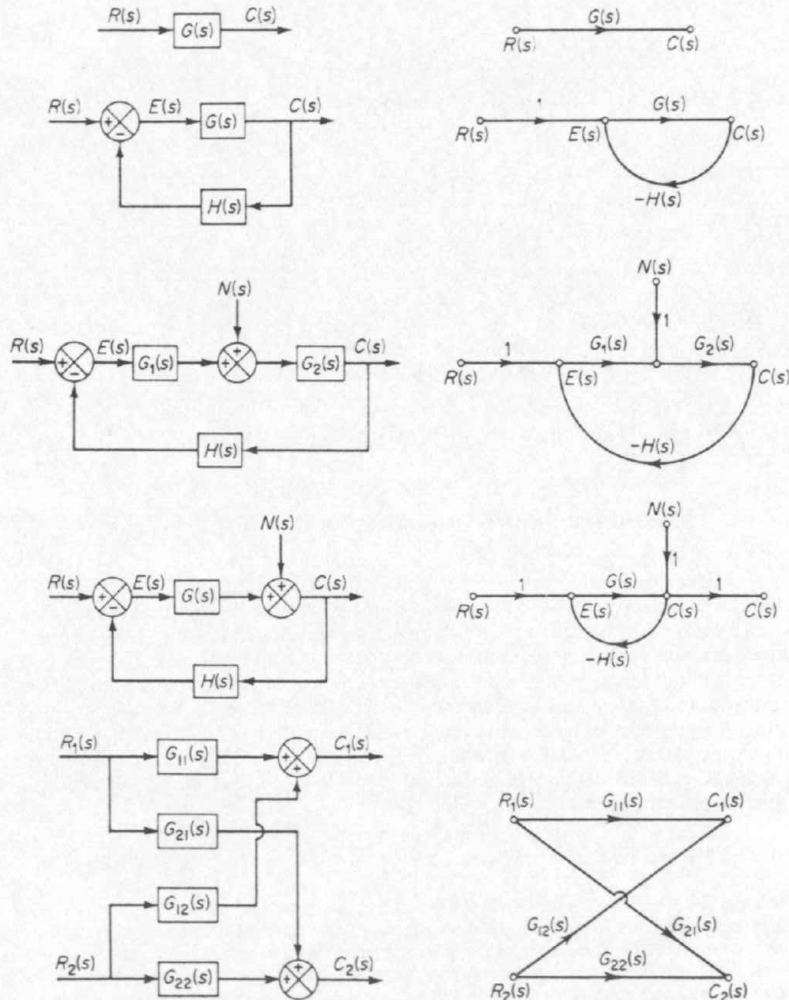


Fig. 4.39 Diagramas de blocos e correspondentes gráficos de fluxo de sinal.

$\sum_{b,c} L_b L_c$ = soma dos produtos de ganhos de todas as combinações possíveis de dois laços que não se tocam

$\sum_{d,e,f} L_d L_e L_f$ = soma dos produtos de ganhos de todas as combinações possíveis de três laços que não se tocam

Δ_k = cofator do k-ésimo determinante do caminho direto do gráfico com os laços tocando o k-ésimo caminho direto removido

(Note que as somas são calculadas com todos os caminhos possíveis desde a entrada até a saída.)

A seguir, ilustraremos a fórmula de ganho de Mason no uso em dois exemplos.

Exemplo 4.4 Considere o sistema mostrado na Fig. 4.40. O gráfico de fluxo de sinal para este sistema é indicado na Fig. 4.41. Vamos obter a função de transferência em malha-fechada $C(s)/R(s)$ usando a fórmula de ganho de Mason.

Neste sistema há apenas um caminho direto entre a entrada $R(s)$ e a saída $C(s)$. O ganho do caminho direto é

$$P_1 = G_1 G_2 G_3$$

Da Fig. 4.41, verificamos que há três laços individuais. Os ganhos destes laços são

$$L_1 = G_1 G_2 H_1$$
$$L_2 = -G_2 G_3 H_2$$
$$L_3 = -G_1 G_2 G_3$$

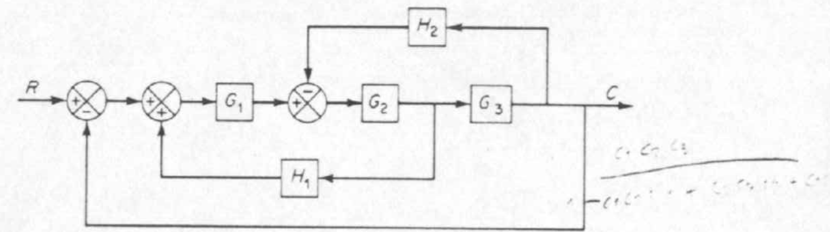


Fig. 4.40 Sistema de laços múltiplos.

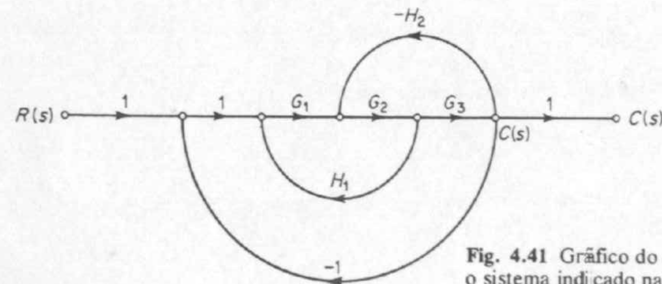


Fig. 4.41 Gráfico do fluxo de sinal para o sistema indicado na Fig. 4.40.

Note que devido ao fato de os três laços possuírem um ramo comum, não há laços que não se tocam. Conseqüentemente, o determinante Δ é dado por

$$\begin{aligned}\Delta &= 1 - (L_1 + L_2 + L_3)\\ &= 1 - G_1G_2H_1 + G_2G_3H_2 + G_1G_2G_3\end{aligned}$$

O cofator Δ_1 do determinante ao longo do caminho direto ligando o nó de entrada e o nó de saída é obtido pela remoção dos laços que tocam este caminho. Desde que o caminho P_1 toca todos os três laços, obtemos

$$\Delta_1 = 1$$

Portanto, o ganho global, entre a entrada $R(s)$ e a saída $C(s)$, ou a função de transferência em malha-fechada, é dado por

$$\begin{aligned}\frac{C(s)}{R(s)} &= P = \frac{P_1\,\Delta_1}{\Delta}\\ &= \frac{G_1G_2G_3}{1 - G_1G_2H_1 + G_2G_3H_2 + G_1G_2G_3}\end{aligned}$$

que é a mesma expressão da função de transferência em malha-fechada obtida pela redução do diagrama de blocos. A fórmula do ganho de Mason fornece o ganho global $C(s)/R(s)$ sem uma redução do gráfico.

Exemplo 4.5 Considere o sistema indicado na Fig. 4.42. Obtenha a função de transferência em malha-fechada $C(s)/R(s)$ usando a fórmula de ganho de Mason.

Neste sistema há três caminhos diretos entre a entrada $R(s)$ e a saída $C(s)$. Os ganhos dos caminhos diretos são

$$\begin{aligned}P_1 &= G_1G_2G_3G_4G_5\\ P_2 &= G_1G_6G_4G_5\\ P_3 &= G_1G_2G_7\end{aligned}$$

Há quatro laços individuais. Os ganhos destes laços são

$$\begin{aligned}L_1 &= -G_4H_1\\ L_2 &= -G_2G_7H_2\\ L_3 &= -G_6G_4G_5H_2\\ L_4 &= -G_2G_3G_4G_5H_2\end{aligned}$$

O laço L_1 não toca o laço L_2. Conseqüentemente, o determinante Δ é dado por

$$\Delta = 1 - (L_1 + L_2 + L_3 + L_4) + L_1L_2 \tag{4.75}$$

O cofator Δ_1 é obtido de Δ removendo os laços que tocam o caminho P_1. Portanto, removendo L_1, L_2, L_3, L_4 e L_1L_2 da Eq. (4.75), obtemos

$$\Delta_1 = 1$$

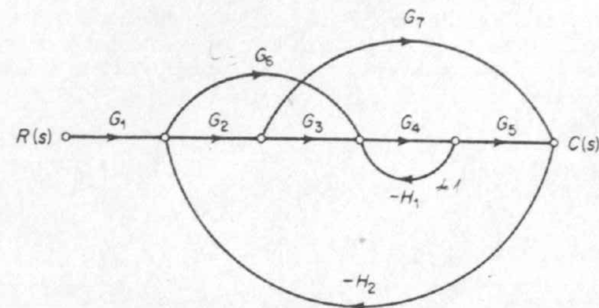


Fig. 4.42 Gráfico do fluxo de sinal para um sistema.

Analogamente, o cofator Δ_2 é

$$\Delta_2 = 1$$

O cofator Δ_3 é obtido removendo-se L_2, L_3, L_4 e L_1L_2 da Eq. (4.75), resultando

$$\Delta_3 = 1 - L_1$$

A função de transferência em malha-fechada $C(s)/R(s)$ é então

$$\begin{aligned}\frac{C(s)}{R(s)} &= P = \frac{1}{\Delta}(P_1\,\Delta_1 + P_2\,\Delta_2 + P_3\,\Delta_3)\\ &= \frac{G_1G_2G_3G_4G_5 + G_1G_6G_4G_5 + G_1G_2G_7(1 + G_4H_1)}{1 + G_4H_1 + G_2G_7H_2 + G_6G_4G_5H_2 + G_4H_1G_2G_7H_2 + G_2G_3G_4G_5H_2}\end{aligned}$$

Comentários conclusivos. A aplicação usual dos gráficos de fluxo de sinal é em diagramação de sistemas. O conjunto de equações descrevendo um sistema linear é representado por um gráfico do fluxo de sinal estabelecendo os nós que representam as variáveis do sistema e interconectando os nós com transmitâncias, diretas ou com pesos, que representam as relações entre as variáveis. Uma fórmula de ganho de Mason pode ser utilizada para estabelecer a relação entre uma entrada e uma saída. (Alternativamente, as variáveis no sistema podem ser eliminadas uma a uma com técnicas de redução.) A fórmula de ganho de Mason é especialmente útil na redução de grandes e complexos diagramas de sistemas em um passo, sem exigir reduções passo-a-passo.

Problemas Típicos e Soluções

Problema A.4.1 A Fig. 4.43 mostra um diagrama esquemático de um acelerômetro. Suponha que a carcaça do acelerômetro está firmemente acoplada a uma estrutura de uma aeronave. (O acelerômetro indica a aceleração de sua carcaça em relação ao espaço inercial.) O ângulo de inclinação θ medido a partir da linha horizontal é suposto constante durante o intervalo de tempo de medida.

Mostre que para entradas em baixas freqüências, a aceleração da carcaça, relativamente ao espaço inercial, pode ser determinada pelo deslocamento da massa m em relação à sua carcaça.

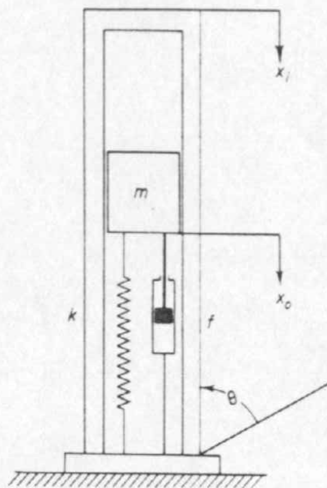


Fig. 4.43 Diagrama esquemático de um acelerômetro.

Solução. Vamos definir

x_i = deslocamento da carcaça em relação ao espaço inercial
x_0 = deslocamento da massa m relativa ao espaço inercial
$y = x_0 - x_i$ = deslocamento da massa m relativamente à carcaça

A equação para este sistema é

$$m\ddot{x}_0 + f(\dot{x}_0 - \dot{x}_i) + k(x_0 - x_i) - mg\,\text{sen}\,\theta = 0$$

Em termos de y, obtemos

$$m\ddot{y} + f\dot{y} + ky = -m\ddot{x}_i + mg\,\text{sen}\,\theta$$

Desde que θ é suposto constante, durante o intervalo de tempo de medida, mg sen θ é constante, e é possível calibrar o deslocamento e definir uma nova variável z tal que

$$z = y - \frac{mg}{k}\,\text{sen}\,\theta$$

Então, obtemos

$$m\ddot{z} + f\dot{z} + kz = -m\ddot{x}_i$$

Se a aceleração de entrada (a aceleração da carcaça relativamente ao espaço inercial) $\ddot{x}_i$ é considerada a entrada do sistema e o deslocamento z é considerado a saída, então a função de transferência do sistema torna-se

$$\frac{Z(s)}{s^2 X_i(s)} = -\frac{1}{s^2 + \dfrac{f}{m}s + \dfrac{k}{m}}$$

Se a freqüência de entrada é muito baixa quando comparada com $\sqrt{k/m}$, então

$$\frac{Z(s)}{s^2 X_i(s)} \doteq -\frac{m}{k}$$

o que significa que $z = y - (mg/k)$ sen θ é aproximadamente proporcional à aceleração da entrada lentamente variada. Portanto, para entradas em baixa freqüência, a aceleração da carcaça relativamente ao espaço inercial pode ser dada por

$$\ddot{x}_i = -\frac{k}{m}\left(y - \frac{mg}{k}\,\text{sen}\,\theta\right)$$

Note que este tipo de acelerômetro possui uma freqüência natural não amortecida suficientemente alta $\sqrt{k/m}$ comparada com a máxima freqüência da entrada a ser medida.

Problema A.4.2 Transmissões com engrenagens são normalmente utilizadas em servomecanismos para reduzir a velocidade, para ampliar o torque ou para obter a mais eficiente transferência de potência casando o elemento excitador com uma dada carga.

Considere o sistema por transmissão com engrenagens mostrado na Fig. 4.44. Neste sistema, uma carga é alimentada por um motor através de uma transmissão por engrenagens. Supondo que a dureza do eixo da transmissão por engrenagem é infinita (e ainda não haja folga nem deformação elástica) e que o número de dentes em cada engrenagem é proporcional ao raio da engrenagem, obtenha o momento de inércia equivalente e a fricção equivalente referidos ao eixo do motor e referidos ao eixo da carga.

Na Fig. 4.44, os números de dentes nas engrenagens 1, 2, 3 e 4 são N_1, N_2, N_3 e N_4, respectivamente. Os deslocamentos angulares dos eixos 1, 2 e 3 são θ_1, θ_2 e θ_3, respectivamente. Portanto, $\theta_2/\theta_1 = N_1/N_2$ e $\theta_3/\theta_2 = N_3/N_4$. O momento de inércia e a fricção viscosa de cada componente da transmissão por engrenagem são denotados por J_1, f_1; J_2, f_2; J_3, f_3, respectivamente. (J_3 e f_3 incluem o momento de inércia e a fricção da carga.)

Solução. Para este sistema de transmissão por engrenagem, podemos obter as três equações seguintes: Para o primeiro eixo,

$$J_1\ddot{\theta}_1 + f_1\dot{\theta}_1 + T_1 = T_m \qquad (4.76)$$

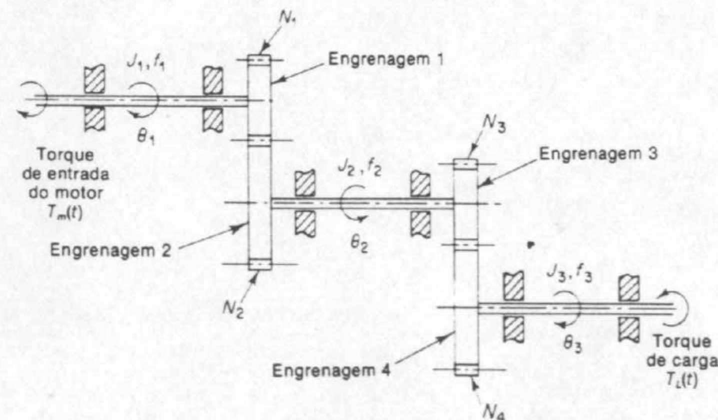


Fig. 4.44 Sistema de transmissão por engrenagens.

onde T_m é o torque desenvolvido pelo motor e T_1 é o torque da carga na engrenagem 1 devido ao resto da transmissão por engrenagem. Para o segundo eixo.

$$J_2\ddot{\theta}_2 + f_2\dot{\theta}_2 + T_3 = T_2 \tag{4.77}$$

onde T_2 é o torque transmitido para a engrenagem 2 e T_3 é o torque de carga na engrenagem 3 devido ao resto da transmissão por engrenagem. Desde que o trabalho efetuado pela engrenagem 1 é igual àquele da engrenagem 2,

$$T_1\theta_1 = T_2\theta_2 \quad \text{ou} \quad T_2 = T_1\frac{N_2}{N_1}$$

Se $N_1/N_2 < 1$, a relação de engrenagem reduz a velocidade bem como amplia o torque. Para o terceiro eixo,

$$J_3\ddot{\theta}_3 + f_3\dot{\theta}_3 + T_L = T_4 \tag{4.78}$$

onde T_L é torque de carga e T_4 é o torque transmitido para a engrenagem 4. T_3 e T_4 são relacionados por

$$T_4 = T_3\frac{N_4}{N_3}$$

e θ_3 e θ_1 são relacionados por

$$\theta_3 = \theta_2\frac{N_3}{N_4} = \theta_1\frac{N_1}{N_2}\frac{N_3}{N_4}$$

Eliminando-se T_1, T_2, T_3 e T_4 das Eqs. (4.76), (4.77) e (4.78), resulta

$$J_1\ddot{\theta}_1 + f_1\dot{\theta}_1 + \frac{N_1}{N_2}(J_2\ddot{\theta}_2 + f_2\dot{\theta}_2) + \frac{N_1N_3}{N_2N_4}(J_3\ddot{\theta}_3 + f_3\dot{\theta}_3 + T_L) = T_m \tag{4.79}$$

Eliminando-se θ_2 e θ_3 da Eq. (4.79) e escrevendo a equação em termos de θ_1 e suas derivadas temporais, obtemos

$$\left[J_1 + \left(\frac{N_1}{N_2}\right)^2 J_2 + \left(\frac{N_1}{N_2}\right)^2\left(\frac{N_3}{N_4}\right)^2 J_3\right]\ddot{\theta}_1 + \left[f_1 + \left(\frac{N_1}{N_2}\right)^2 f_2 + \left(\frac{N_1}{N_2}\right)^2\left(\frac{N_3}{N_4}\right)^2 f_3\right]\dot{\theta}_1 + \left(\frac{N_1}{N_2}\right)\left(\frac{N_3}{N_4}\right)T_L = T_m \tag{4.80}$$

Portanto, o momento de inércia e a fricção equivalentes da transmissão por engrenagem referidos ao eixo 1 são dados por

$$J_{1eq} = J_1 + \left(\frac{N_1}{N_2}\right)^2 J_2 + \left(\frac{N_1}{N_2}\right)^2\left(\frac{N_3}{N_4}\right)^2 J_3$$

$$f_{1eq} = f_1 + \left(\frac{N_1}{N_2}\right)^2 f_2 + \left(\frac{N_1}{N_2}\right)^2\left(\frac{N_3}{N_4}\right)^2 f_3$$

Analogamente, o momento de inércia e a fricção equivalentes da transmissão por engrenagem referidos ao eixo da carga são

$$J_{3eq} = J_3 + \left(\frac{N_4}{N_3}\right)^2 J_2 + \left(\frac{N_2}{N_1}\right)^2\left(\frac{N_4}{N_3}\right)^2 J_1$$

$$f_{3eq} = f_3 + \left(\frac{N_4}{N_3}\right)^2 f_2 + \left(\frac{N_2}{N_1}\right)^2\left(\frac{N_4}{N_3}\right)^2 f_1$$

A relação entre J_{1eq} e J_{3eq} é portanto

$$J_{1eq} = \left(\frac{N_1}{N_2}\right)^2\left(\frac{N_3}{N_4}\right)^2 J_{3eq}$$

e a relação entre f_{1eq} e f_{3eq} é

$$f_{1eq} = \left(\frac{N_1}{N_2}\right)^2\left(\frac{N_3}{N_4}\right)^2 f_{3eq}$$

O efeito de J_2 e J_3 no momento de inércia equivalente é determinado pelas relações de engrenagem N_1/N_2 e N_3/N_4. Para transmissões por engrenagem redutoras de velocidade as relações N_1/N_2 e N_3/N_4 são normalmente menores do que a unidade. Se $N_1/N_2 << 1$ e $N_3/N_4 << 1$, então os efeitos de J_2 e J_3 no momento de inércia equivalente J_{1eq} são desprezíveis. Comentários análogos aplicam-se à fricção equivalente f_{1eq} da transmissão por engrenagem. Em termos do momento de inércia equivalente J_{1eq} e da fricção equivalente f_{1eq}, a Eq. (4.80) pode ser simplificada para fornecer

$$J_{1eq}\ddot{\theta}_1 + f_{1eq}\dot{\theta}_1 + nT_L = T_m$$

onde

$$n = \frac{N_1}{N_2}\frac{N_3}{N_4}$$

Considerando $T_m - nT_L$ como a entrada para o sistema e θ_1 como a saída, a função de transferência deste sistema é

$$\frac{\Theta_1(s)}{T_m(s) - nT_L(s)} = \frac{1}{s(J_{1eq}s + f_{1eq})}$$

O diagrama de blocos é mostrado na Fig. 4.45(a). Se $(T_m - nT_L)$ e θ_3 são consideradas como a entrada e a saída, respectivamente, então a função de transferência torna-se

$$\frac{\Theta_3(s)}{T_m(s) - nT_L(s)} = \frac{\frac{1}{n}}{s(J_{3eq}s + f_{3eq})}$$

Desde que

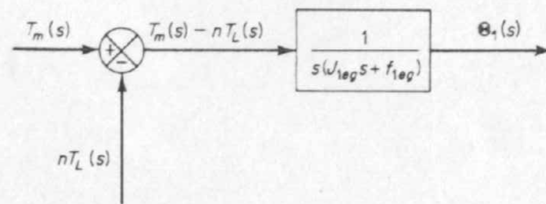
$$\frac{\Theta_3(s)}{\Theta_1(s)} = \frac{N_1N_3}{N_2N_4} = n$$

em termos de J_{1eq} e f_{1eq}, sua função de transferência pode ser escrita

$$\frac{\Theta_3(s)}{T_m(s) - nT_L(s)} = \frac{n}{s(J_{1eq}s + f_{1eq})}$$

O diagrama de blocos para este caso é mostrado na Fig. 4.45(b).

Da análise acima, pode ser visto que em um sistema onde um servomotor está excitando uma carga com inércia e uma carga com fricção viscosa, a inércia e fricção da carga são, para o eixo do motor, os valores originais multiplicados pelo quadrado da relação da engrenagem. Se



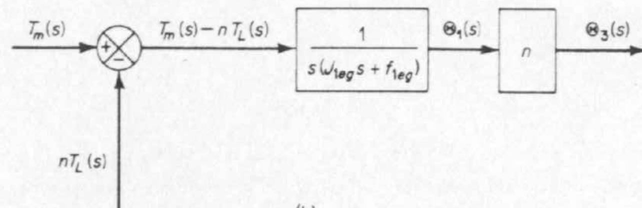


Fig. 4.45 Diagramas de blocos no sistema de transmissão por engrenagens mostrado na Fig. 4.44.

a relação da engrenagem for um número pequeno, o momento de inércia e a fricção do servomotor possuem efeitos dominantes no comportamento dinâmico do sistema.

Problema A.4.3 Mostre que as relações torque-inércia referidas ao eixo do motor e ao eixo da carga diferem uma da outra por um fator n. Mostre também que as relações torque ao quadrado-inércia referidas ao eixo do motor e ao eixo da carga são as mesmas.

Solução. Suponha que $T_{\text{máx.}}$ é o torque máximo que pode ser produzido no eixo do motor. Então a relação torque-inércia referida ao eixo do motor é

$$\frac{T_{\text{máx.}}}{J_m + n^2 J_L}$$

onde

J_m = momento de inércia do rotor
J_L = momento de inércia da carga
n = relação de engrenagem

A relação torque-inércia referida ao eixo da carga é

$$\frac{\dfrac{T_{\text{máx.}}}{n}}{J_L + \dfrac{J_m}{n^2}}$$

Obviamente, elas diferem por um fator n. Portanto, na comparação das relações torque-inércia de motores, devemos necessariamente especificar qual eixo é a referência.

Note que a relação torque ao quadrado-inércia referida ao eixo do motor é

$$\frac{T^2_{\text{máx.}}}{J_m + n^2 J_L}$$

e aquela referida ao eixo da carga é

$$\frac{\dfrac{T^2_{\text{máx.}}}{n^2}}{J_L + \dfrac{J_m}{n^2}}$$

Estas duas relações são evidentemente as mesmas.

Problema A.4.4 O momento de inércia J de um motor pode ser disponível no catálogo do fabricante ou pode ser calculado se as dimensões e o material do rotor forem conhecidos.

O momento de inércia efetivo de um sistema motor e transmissão por engrenagens também pode ser determinado por métodos experimentais. Discuta os métodos para determinar o momento de inércia J de um motor c.a. bifásico.

Solução. Determinar o momento de inércia efetivo J de um sistema com um motor de transmissão por engrenagens é possível aplicando-se uma entrada em degrau e_i e registrando-se a resposta transitória. A função de transferência de um motor c.a. bifásico é dada pela Eq. (4.35) como segue:

$$\frac{\Theta(s)}{E_i(s)} = \frac{K_c}{Js^2 + (f + K_n)s}$$

Se a tensão de entrada e_i é uma entrada em degrau de altura E_1, então a velocidade de eixo resultante é

$$s\Theta(s) = \frac{K_c E_1}{s(Js + f + K_n)}$$

Portanto a velocidade angular $\omega(t)$ é

$$\omega(t) = \dot{\theta}(t) = \frac{K_c E_1}{f + K_n}\left[1 - \exp\left(-\frac{f + K_n}{J}t\right)\right]$$

Do registro de $\omega(t)$ *versus* tempo, a constante de tempo $J/(f + K_n)$ pode ser determinada.

Desde que a velocidade angular em regime estacionário é

$$\omega(\infty) = \frac{K_c E_1}{f + K_n}$$

obtemos

$$f + K_n = \frac{K_c E_1}{\omega(\infty)}$$

Para determinar $f + K_n$, necessitamos conhecer K_c. [E_1 e $\omega(\infty)$ são grandezas conhecidas.] O

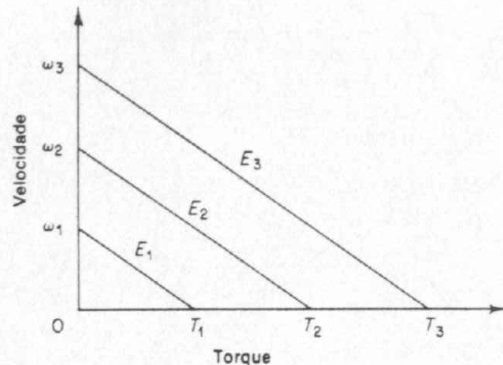


Fig. 4.46 Curvas torque-velocidade.

valor da constante de torque K_c pode ser determinado a partir das curvas torque-velocidade, como mostrado na Fig. 4.46, desde que estas curvas são paralelas. Das curvas obtemos

$$K_c = \frac{T_1}{E_1} = \frac{T_2}{E_2} = \frac{T_3}{E_3}$$

Uma vez conhecida a constante de tempo $J/(f + K_n)$, a velocidade angular em regime estacionário $\omega(\infty)$, e a constante de torque K_c, pode ser determinado o momento de inércia J.

O momento de inércia J do rotor também pode ser determinado experimentalmente por meio de experiência com o pêndulo de torção. Se o coeficiente de amortecimento é muito pequeno, o período de oscilação do pêndulo é relacionado com J pela seguinte equação:

$$J = \frac{kT^2}{4\pi^2}$$

onde

J = momento de inércia do rotor
k = constante da mola de torção do eixo elástico no qual o rotor do motor é acoplado
T = período da oscilação

Problema A.4.5 Obtenha a função de transferência do servomotor bifásico cuja curva torque-velocidade é indicada na Fig. 4.47. As máximas tensões nominais de fase fixa e de fase de controle são 115 volts. O momento de inércia J do rotor (incluindo o efeito da carga) é $7{,}77 \times 10^{-4}$ oz-in-s², e o coeficiente de fricção-viscosa do motor (incluindo o efeito da carga) é 0,005 oz-in/rad/s.

Solução. A equação para a curva torque-velocidade é

$$T + K_n\dot{\theta} = K_c E_c$$

onde

$$K_n = \frac{5}{4000}\frac{60}{2\pi} = 0{,}0119 \text{ oz-in /rad/s}$$

$$K_c = \frac{5}{115} = 0{,}0435 \text{ oz-in /volt}$$

Portanto,

$$T + 0{,}0119\dot{\theta} = 0{,}0435 E_c$$

A constante de ganho do motor K_m é então

$$K_m = \frac{K_c}{f + K_n} = \frac{0{,}0435}{0{,}005 + 0{,}0119} = 2{,}57$$

A constante de tempo do motor T_m é

$$T_m = \frac{J}{f + K_n} = \frac{7{,}77 \times 10^{-4}}{5 \times 10^{-3} + 11{,}9 \times 10^{-3}} = 0{,}046$$

Usando os valores numéricos obtidos, a função de transferência dada pela Eq. (4.35) é a seguinte:

$$\frac{\Theta(s)}{E_c(s)} = \frac{K_m}{s(T_m s + 1)} = \frac{2{,}57}{s(0{,}046s + 1)}$$

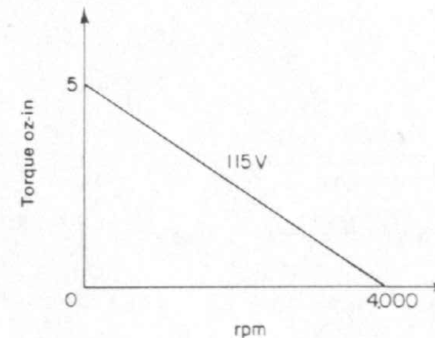


Fig. 4.47 Curva torque-velocidade.

Problema A.4.6 Supondo que um servomotor bifásico possui uma curva linear torque-velocidade de modo que a velocidade sem carga é ω_0 e o torque completamente freado é T_s, determine a máxima potência de saída no eixo do motor $P_{máx}$.

Solução. A curva torque-velocidade é

$$T = T_s - \left(\frac{\omega}{\omega_0}\right)T_s$$

A potência de saída do eixo P é dada por

$$P = T\omega$$
$$= \left(T_s - \frac{\omega}{\omega_0}T_s\right)\omega$$

Para determinar $P_{máx}$, vamos diferenciar P em relação a ω:

$$\frac{dP}{d\omega} = T_s - 2\frac{\omega}{\omega_0}T_s$$

Impondo que $dP/d\omega = 0$, obtemos

$$\omega = \frac{\omega_0}{2}$$

Claramente, $d^2P/d\omega^2 = -2T_s/\omega_0 < 0$. Conseqüentemente P é um máximo em $\omega = \omega_0/2$. $P_{máx}$ é

$$P_{máx} = T\omega\,|_{\omega=\omega_0/2}$$
$$= \left(T_s - \frac{T_s}{2}\right)\frac{\omega_0}{2}$$
$$= \frac{T_s\omega_0}{4}$$

Para o servomotor considerado no Problema A.4.5, $T_s = 5$ oz-in e $\omega_0 = 4.000$ rpm. A potência de saída máxima $P_{máx}$ é

$$P_{máx} = \frac{5 \times 4000}{4} \times \frac{1}{1352}$$
$$= 3{,}7 \text{ watts}$$

A potência máxima ocorre em $\omega = 2.000$ rpm.

Problema A.4.7 A máxima aceleração atingível com o servomotor pode ser indicada pela relação torque-inércia, sendo a relação do máximo torque, na condição de freagem completa e total do eixo do motor, e a inércia do rotor. Quanto maior esta relação, melhor será a característica de aceleração. Referindo-se ao servomotor bifásico considerado no Problema A.4.5, determine a relação torque-inércia.

Solução. O momento de inércia J do rotor do servomotor bifásico é $7{,}77 \times 10^{-4}$ oz-in-s². O torque máximo na condição de freagem completa é 5 oz-in. Conseqüentemente a relação torque-inércia é dada por

$$\frac{T_{máx.}}{J} = \frac{5}{7{,}77 \times 10^{-4}}$$
$$= 6{,}435 \text{ rad/s}^2$$

Na seleção de um servomotor, é importante satisfazer um dado requisito em relação à aceleração.

Problema A.4.8 Considere o motor c.c. controlado por armadura indicado na Fig. 4.48. Na Seção 4.5 foi demonstrado que

$$e_b = K_b\dot{\theta}$$
$$T = K_f i_f K_1 i_a = K i_a$$

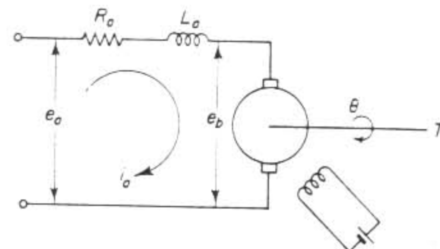


Fig. 4.48 Motor c.c. controlado por armadura.

onde K_b é a constante de força-contra-eletromotriz, i_f é a corrente de campo constante e K_f e K_1 são constantes. Demonstre que

$$\frac{K_b}{K} = 1{,}356 \text{ watts-s/ft-lb}$$

Solução. A potência mecânica $T\dot{\theta}$ deve ser igual à potência elétrica $e_b i_a$ desenvolvida pela corrente de armadura fluindo através da força-contra-eletromotriz da armadura. Então, de

$$e_b i_a = K_b\dot{\theta}i_a \text{ (watts)} = \frac{K_b\dot{\theta}i_a}{746} \text{ (hp)}$$

$$T\dot{\theta} = K i_a\dot{\theta} \text{ (ft-lb/s)} = \frac{K i_a\dot{\theta}}{550} \text{ (hp)}$$

obtemos

$$\frac{K_b}{K} = 1{,}356 \text{ watts-s/ft-lb}$$

Problema A.4.9 Referindo-se ao motor c.c. mostrado na Fig. 4.48, para valores fixados da tensão de armadura, as curvas de torque *versus* velocidade podem ser representadas por

$$T = \frac{K}{R_a}(e_a - K_b\omega)$$

Determine K e K_b das curvas torque *versus* velocidade.

Solução. Para a condição em vazio (sem carga), $T = 0$. Portanto,

$$K_b = \frac{e_a}{\omega_0}$$

onde

ω_0 = velocidade sem carga

Se o motor for completamente freado, então $\omega = 0$, e obtemos

$$K = \frac{T_s R_a}{e_a}$$

onde

T_s = torque com freagem total.

Problema A.4.10 Determine a função de transferência $\Theta(s)/E_a(s)$ do motor c.c. controlado por armadura mostrado na Fig. 4.48. No diagrama

e_a = 26 volts
L_a = desprezível
$J = 2,59 \times 10^{-4}$ oz-in-s²
$f = 3 \times 10^{-3}$ oz-in/rad/s

Suponha que o torque com freagem completa é 10 oz-in e a velocidade sem carga é 520 rad/s.

Solução. Da Eq. (4.43), a função de transferência $\Theta(s)/E_a(s)$ é

$$\frac{\Theta(s)}{E_a(s)} = \frac{K_m}{s(T_m s + 1)}$$
$$= \frac{K}{s(R_a J s + R_a f + K K_b)}$$

onde

$$K = \frac{T_s R_a}{e_a} \qquad (T_s = \text{torque freado})$$

$$K_b = \frac{e_a}{\omega_0} \qquad (\omega_0 = \text{velocidade sem carga})$$

No Problema A.4.8, verificamos que

$$\frac{K_b}{K} = 1,356 \text{ watts-s/ft-lb}$$
$$= 7,06 \times 10^{-3} \text{ watts-s/in-oz}$$

Portanto,

$$K_b = \frac{26}{520} = 5 \times 10^{-2} \text{ volt/rad/s}$$

e

$$K = \frac{K_b}{7,06 \times 10^{-3}}$$
$$= \frac{5 \times 10^{-2}}{7,06 \times 10^{-3}}$$
$$= 7,08 \text{ in-oz/amp}$$

O valor de R_a é então obtido como

$$R_a = \frac{K e_a}{T_s}$$
$$= 7,08 \times \frac{26}{10}$$
$$= 18,4 \text{ ohms}$$

A função de transferência $\Theta(s)/E_a(s)$ é então

$$\frac{\Theta(s)}{E_a(s)} = \frac{7,08}{s[(18,4)(2,59 \times 10^{-4})s + (18,4)(3 \times 10^{-3}) + (7,08)(5 \times 10^{-2})]}$$
$$= \frac{17,3}{s(0,0116s + 1)}$$

Problema A.4.11 Determine a função de transferência $\Theta(s)/E_f(s)$ do motor c.c. controlado por campo mostrado na Fig. 4.26(a). Supor que no diagrama

e_f = 110 volts
i_a = 15 ampères
$\omega_{ss} = \dot{\theta}_{ss}$ = 1.200 rpm

L_f = 20 henrys
R_f = 120 ohms
J = 1 lb-ft-s²
f = 0.5 lb-ft/rad/s

Solução. Da Eq. (4.52), a função de transferência $\Theta(s)/E_f(s)$ é

$$\frac{\Theta(s)}{E_f(s)} = \frac{K_2}{s(L_f s + R_f)(Js + f)}$$

Desta equação, obtemos a velocidade angular em regime estacionário ω_{ss} como segue: Note que para uma entrada em degrau de e_f = 110 volts, obtemos

$$\Theta(s) = \frac{K_2}{s(L_f s + R_f)(Js + f)} \frac{110}{s}$$

Portanto,

$$\omega_{ss} = \lim_{s\to 0} s[s\Theta(s)]$$
$$= \lim_{s\to 0} \frac{110K_2}{(L_f s + R_f)(Js + f)}$$
$$= \frac{110K_2}{R_f f}$$

Desde que ω_{ss} = 1.200 rpm = 20 × 2π rad/s, R_f = 120 ohms, f = 0,5 lb-ft/rad/s, obtemos

$$40\pi = \frac{110K_2}{120 \times 0,5}$$

ou

$$K_2 = \frac{40\pi \times 60}{110} \text{ ohms lb-ft/volt}$$
$$= 68,5 \text{ lb-ft/amp}$$

A função de transferência $\Theta(s)/E_f(s)$ é então obtida como segue:

$$\frac{\Theta(s)}{E_f(s)} = \frac{68,5}{s(20s + 120)(s + 0,5)}$$
$$= \frac{1,14}{s(0,167s + 1)(2s + 1)}$$

Problema A.4.12 Considere o servomecanismo de posição indicado na Fig. 4.49. Suponha que a entrada do sistema seja a posição do eixo de referência e a saída do sistema seja a posição do eixo de saída. Suponha os seguintes valores numéricos para as constantes do sistema:

r = deslocamento angular do eixo de entrada de referência, radianos
c = deslocamento angular do eixo de saída, radianos
θ = deslocamento angular do eixo do motor, radianos
K_1 = ganho do detector de erro do potenciômetro = 24/π volts/rad
K_p = ganho do amplificador = 10 volts/volt
R_f = resistência do enrolamento de campo = 2 ohms
L_f = indutância do enrolamento de campo = 0,1 henry
i_f = corrente de enrolamento de campo, ampères
e_f = tensão de campo aplicada, volts
K_2 = constante de torque do motor = 0,05 Newton-m/ampères
n = relação de engrenagem = 1/10
J = momento equivalente de inércia do motor e carga referido ao eixo do motor = 0,02 kg-m²
f = coeficiente de fricção-viscosa equivalente do motor e carga referente ao eixo do motor = 0,02 Newton-m/rad/s

Desenhe um diagrama de blocos do sistema. Determine a função de transferência de cada bloco. Finalmente, simplifique o diagrama de blocos.

Solução. As equações que descrevem a dinâmica do sistema são as seguintes:

Para o detector de erro potenciométrico:

$$E(s) = K_1[R(s) - C(s)] = \frac{24}{\pi}[R(s) - C(s)] \tag{4.81}$$

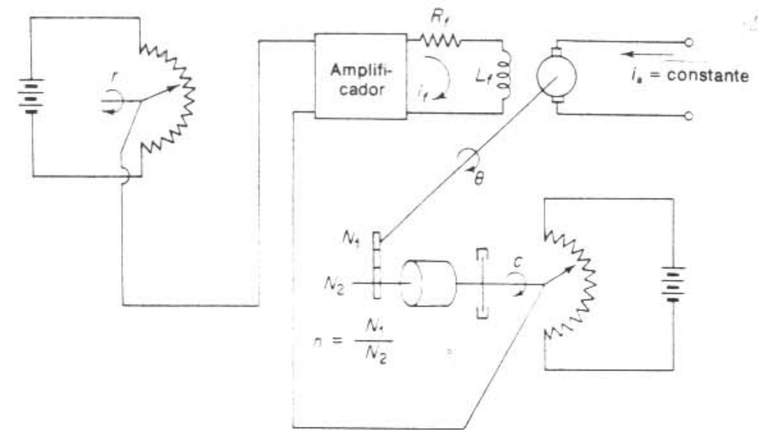


Fig. 4.49 Servomecanismo de posição.

Para o amplificador:

$$E_f(s) = K_p E(s) = 10E(s) \tag{4.82}$$

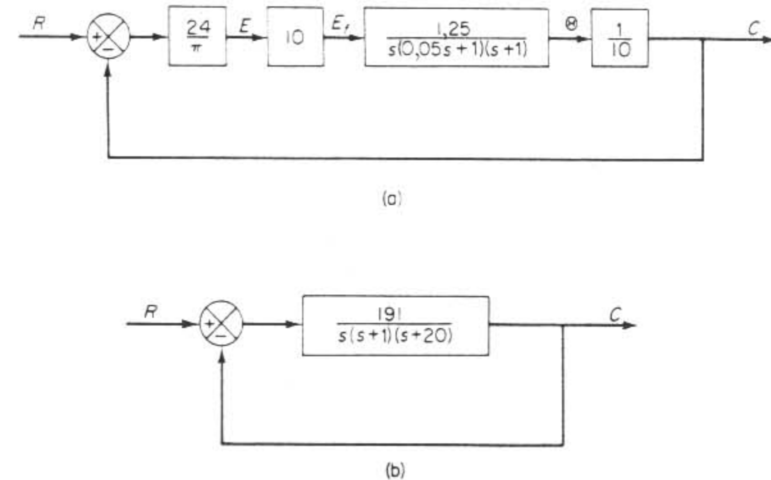


Fig. 4.50 (a) Diagrama de blocos do servomecanismo de posição indicado na Fig. 4.49; (b) diagramas de blocos simplificado.

Para o motor controlado por campo: Da Eq. (4.52), verificamos que

$$\frac{\Theta(s)}{E_f(s)} = \frac{K_m}{s(T_f s + 1)(T_m s + 1)}$$

onde

$$K_m = \frac{K_2}{R_f f} = \frac{0{,}05}{2 \times 0{,}02} = 1{,}25 \text{ rad volt-s}$$

$$T_f = \frac{L_f}{R_f} = \frac{0{,}1}{2} = 0{,}05 \text{ s}$$

$$T_m = \frac{J}{f} = \frac{0{,}02}{0{,}02} = 1 \text{ s}$$

Portanto,

$$\frac{\Theta(s)}{E_f(s)} = \frac{1{,}25}{s(0{,}05s + 1)(s + 1)} \tag{4.83}$$

Das Eqs. (4.81), (4.82) e (4.83), obtemos o diagrama de blocos indicado na Fig. 4.50(a). Uma simplificação do diagrama de blocos da Fig. 4.50(a) resulta na Fig. 4.50(b).

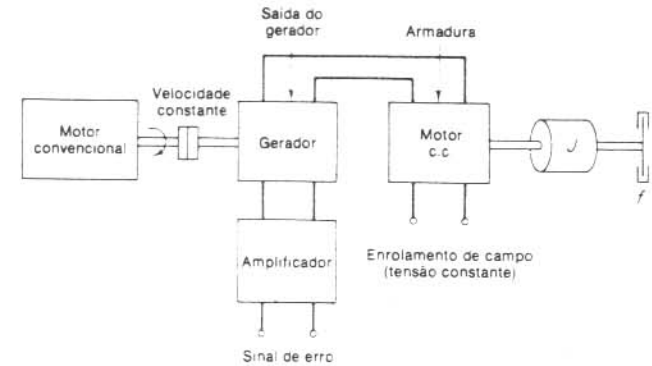
Problema A.4.13 Nos sistemas de controle realimentados são utilizados geradores c.c. a fim de produzir-se uma grande amplificação de potência. (Amplificadores a válvula, amplificadores transistorizados etc. não são convenientes para utilização no controle direto de um motor c.c. com mais de 100 watts, se forem desejadas pequenas constantes de tempo e boa linearidade.) No diagrama mostrado na Fig. 4.51(a), o gerador c.c. excita um motor c.c. controlado por armadura. No diagrama esquemático da Fig. 4.51(b),

R_f = resistência do enrolamento de campo do gerador c.c., ohms
L_f = indutância do enrolamento de campo do gerador c.c., henrys
i_g = corrente do enrolamento de campo do gerador c.c., amperes
e_f = tensão de enrolamento de campo aplicada do gerador c.c., volts
R_a = resistência do enrolamento de armadura, ohms
L_a = indutância do enrolamento de armadura, henrys
i_a = corrente do enrolamento de armadura, amperes
e_a = tensão gerada na armadura, volts
$\dot{\theta}_g$ = velocidade constante do gerador c.c., rad/s
i_f = corrente de campo do motor c.c. (constante), amperes
e_b = força-contra-eletromotriz, volts
θ = deslocamento angular do eixo do motor, radianos
T = torque fornecido pelo motor, N-m
J = momento de inércia do motor e carga referidos ao eixo do motor, kg-m²
f = coeficiente de fricção-viscosa do motor e da carga referidos ao eixo do motor, N-m/rad/s

O gerador é girado com uma velocidade constante $\dot{\theta}_g$. A tensão gerada na armadura e_a é determinada pela tensão e_f aplicada ao campo do gerador. Deduza a função de transferência entre θ e e_f.

Solução. As equações para este sistema são

$$L_f \frac{di_g}{dt} + R_f i_g = e_f \tag{4.84}$$



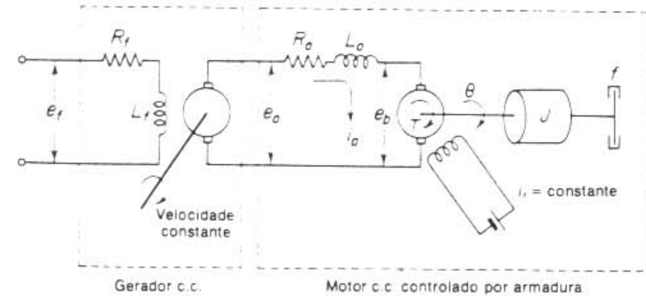


Fig. 4.51 (a) Um sistema gerador-motor c.c.; (b) diagrama esquemático.

e

$$e_a = K\dot{\theta}_g \psi_g$$

onde K é uma constante e ψ_g é o fluxo no entreferro de ar. Desde que a corrente no enrolamento de campo e o fluxo no entreferro de ar são proporcionais e $\dot{\theta}_g$ é uma constante, e_a pode ser escrito

$$e_a = K_a i_g \tag{4.85}$$

onde K_a é uma constante. Eliminando i_g das Eqs. (4.84) e (4.85), obtemos

$$L_f \frac{de_a}{dt} + R_f e_a = K_a e_f$$

A função de transferência entre a tensão gerada na armadura e_a e a tensão no enrolamento e_f é então

$$\frac{E_a(s)}{E_f(s)} = \frac{K_a}{L_f s + R_f} \tag{4.86}$$

No motor c.c. controlado por armadura, a armadura é energizada pela saída do gerador e a corrente de campo é mantida constante. A função de transferência entre θ e e_a foi obtida na Seção 4.5 e resultou na Eq. (4.42). Reescrevendo-a,

$$\frac{\Theta(s)}{E_a(s)} = \frac{K}{s[L_a J s^2 + (L_a f + R_a J)s + R_a f + KK_b]} \tag{4.87}$$

Portanto, a função de transferência entre θ e e_f é obtida a partir das Eqs. (4.86) e (4.87) como

$$\frac{\Theta(s)}{E_f(s)} = \frac{KK_a}{(L_f s + R_f)s[L_a J s^2 + (L_a f + R_a J)s + R_a f + KK_b]} \tag{4.88}$$

A indutância L_a no circuito de armadura normalmente é pequena e pode ser desprezada. Se L_a é desprezada, então a função de transferência dada pela Eq. (4.88) reduz-se a

$$\frac{\Theta(s)}{E_f(s)} = \frac{K_g K_m}{s(T_f s + 1)(T_m s + 1)}$$

onde

$K_g = K_a/R_f$ = constante de ganho do gerador

$T_f = L_f/R_f$ = constante de tempo do gerador

$K_m = K/(R_a f + KK_b)$ = constante de ganho do motor

$T_m = R_a J/(R_a f + KK_b)$ = constante de tempo do motor

Problema A.4.14 Considerando pequenos desvios em relação à operação em regime estacionário, desenhe o diagrama de blocos do sistema de aquecimento de ar indicado na Fig. 4.52. Suponha que a perda de calor para o ambiente e a capacitância térmica das partes metálicas do aquecedor são desprezíveis.

Solução. Vamos definir

$\bar{\Theta}_i$ = temperatura em regime estacionário do ar na abertura de entrada, °C
$\bar{\Theta}_0$ = temperatura em regime estacionário do ar na abertura de saída, °C
G = taxa de fluxo de ar através da câmara de aquecimento, kg/s
M = ar contido na câmara de aquecimento, kg
c = calor específico do ar, cal/kg °C
R = resistência térmica, °C s/cal
C = capacitância térmica do ar contido na câmara de aquecimento = Mc, cal/°C
$\bar{H}$ = entrada de calor em regime estacionário, cal/s

Vamos supor que a entrada de calor seja subitamente variada de $\bar{H}$ para $\bar{H} + h$ e a temperatura de ar na abertura de entrada seja subitamente variada de $\bar{\Theta}_i$ para $\bar{\Theta}_i + \theta_i$. Então, a temperatura do ar na abertura de saída será variada desde $\bar{\Theta}_0$ para $\bar{\Theta}_0 + \theta_0$.

A equação que descreve o comportamento do sistema é

$$C\, d\theta_0 = [h + Gc(\theta_i - \theta_0)]\, dt$$

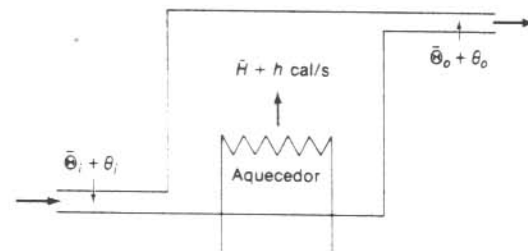


Fig. 4.52 Sistema de aquecimento de ar.

ou

$$C\frac{d\theta_0}{dt} = h + Gc(\theta_i - \theta_0)$$

Notando que

$$Gc = \frac{1}{R}$$

obtemos

$$C\frac{d\theta_0}{dt} = h + \frac{1}{R}(\theta_i - \theta_0)$$

ou

$$RC\frac{d\theta_0}{dt} + \theta_0 = Rh + \theta_i$$

Considerando as transformadas de Laplace de ambos os membros desta última equação e substituindo a condição inicial $\theta_0(0) = 0$, obtemos

$$\Theta_0(s) = \frac{R}{RCs + 1}H(s) + \frac{1}{RCs + 1}\Theta_i(s)$$

O diagrama de blocos do sistema correspondente a esta equação é indicado na Fig. 4.53.

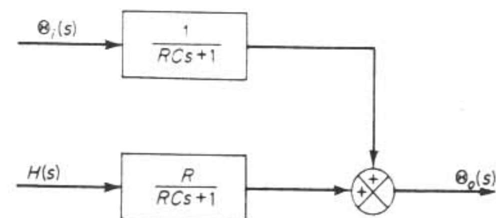


Fig. 4.53 Diagrama de blocos do sistema de aquecimento de ar indicado na Fig. 4.52.

Problema A.4.15 Considere uma variação em um gás ideal desde um estado representado por (p_1, v_1, T_1) até o estado representado por (p_2, v_2, T_2). Se considerarmos a temperatura constante em T porém variando a pressão desde p_1 até p_2, então o volume do gás variará de v_1 para v' de modo que

$$p_1 v_1 = p_2 v' \tag{4.89}$$

Considere agora a pressão constante, porém varie a temperatura para T_2. Então o volume do gás atinge v_2. Portanto,

$$\frac{v'}{T_1} = \frac{v_2}{T_2} \tag{4.90}$$

Eliminando v' entre as Eqs. (4.89) e (4.90), obtemos

$$\frac{p_1 v_1}{T_1} = \frac{p_2 v_2}{T_2}$$

Isto significa que, para uma quantidade fixa de gás, independente de qual variação física ocorra, pv/T será constante. Podemos portanto escrever

$$pv = kT$$

onde o valor da constante k depende da quantidade e natureza do gás considerado.

Tratando-se de sistemas com gás verifica-se ser conveniente trabalhar em quantidades molares desde que 1 mol de qualquer gás contém o mesmo número de moléculas. Conseqüentemente 1 mol ocupa o mesmo volume se medido sob as mesmas condições de pressão e temperaturas.

Se considerarmos 1 mol de um gás, então

$$p\bar{v} = \bar{R}T \tag{4.91}$$

O valor de $\bar{R}$ é o mesmo para todos os gases sob quaisquer condições. A constante $\bar{R}$ é denominada constante universal dos gases. Na condição de pressão e temperatura padrões (isto é, 0°C e 1 atm), 1 kg-mol de qualquer gás ocupa 359 ft³. (Por exemplo, a 0°C e 14,7 psia o volume ocupado por 2 lb de hidrogênio, 32 lb de oxigênio ou 28 lb de nitrogênio é o mesmo, 359 ft³.) Este volume é chamado volume molar e é denotado por $\bar{v}$.

Obtenha o valor da constante universal dos gases.

Solução. Substituindo p = 14,7 lb/in², $\bar{v}$ = 359 ft³/lb-mol e T = 492°R na Eq. (4.91), obtemos

$$\bar{R} = \frac{14{,}7 \times 144 \times 359}{492}$$
$$= 1.545 \text{ ft-lb/mol °R}$$
$$= 1.987 \text{ Btu/lb-mol °R}$$

Problema A.4.16 O peso molecular de uma substância pura é o peso de uma molécula da substância comparado com o peso de um átomo de oxigênio, que é considerado como 16. Isto é, o peso molecular do dióxido de carbono (CO_2) é 12 + (16 × 2) = 44. Os pesos moleculares do oxigênio (molecular) e do vapor de água são 32 e 18, respectivamente.

Obtenha o volume específico de uma mistura que consiste em 100 ft³ de oxigênio, 5 ft³ de dióxido de carbono e 20 ft³ de vapor de água. Suponha que a temperatura e a pressão sejam 70°F e 14,7 psia, respectivamente.

Solução. O peso molecular médio da mistura é

$$M = \left(32 \times \frac{100}{125}\right) + \left(44 \times \frac{5}{125}\right) + \left(18 \times \frac{20}{125}\right) = 25{,}6 + 1{,}76 + 2{,}88 = 30{,}24$$

Portanto,

$$v = \frac{\bar{R}}{M}\frac{T}{p}$$
$$= \frac{1545 \times 530}{30{,}24 \times 14{,}7 \times 144} = 12{,}8 \text{ ft}^3/\text{lb}$$

Problema A.4.17 O valor da constante dos gases para qualquer gás pode ser determinado a partir de observações experimentais precisas dos valores simultâneos de p, v e T.

Obtenha a constante do gás R_{ar} para o ar. Note que a 32°F e 14,7 psia o volume específico do ar é 12,39 ft³/lb. Então obtenha a capacitância de um recipiente de pressão com 20 ft³ que contém ar a 160°F. Supor que o processo de expansão é isotérmico.

Solução

$$R_{ar} = \frac{pv}{T} = \frac{14{,}7 \times 144 \times 12{,}39}{460 + 32} = 53{,}3 \text{ ft-lb/lb °R}$$

Referindo-se à Eq. (4.64), a capacitância de um recipiente de pressão de 20 ft³ é:

$$C = \frac{V}{nR_{ar}T} = \frac{20}{1 \times 53{,}3 \times 620} = 6{,}05 \times 10^{-4}\,\frac{\text{lb}}{\text{lb/ft}^2}$$

Problema A.4.18 Considere o sistema indicado na Fig. 4.54. Obtenha a função de transferência em malha-fechada $H(s)/Q(s)$.

Solução. No sistema dado, há apenas um caminho direto que liga a entrada $Q(s)$ e a saída $H(s)$. Portanto,

$$P_1 = \frac{1}{C_1 s}\frac{1}{R_1}\frac{1}{C_2 s}$$

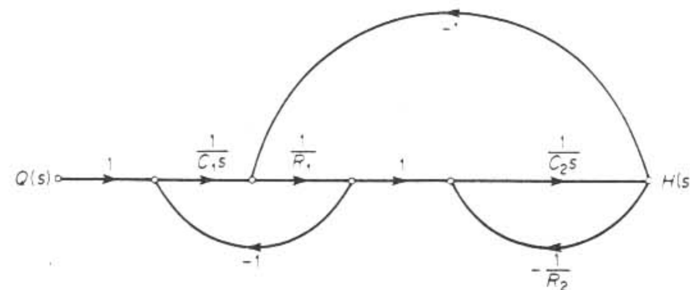


Fig. 4.54 Gráfico de fluxo de sinal de um sistema de controle.

Há três laços individuais. Portanto,

$$L_1 = -\frac{1}{C_1 s}\frac{1}{R_1}$$

$$L_2 = -\frac{1}{C_2 s}\frac{1}{R_2}$$

$$L_3 = -\frac{1}{R_1}\frac{1}{C_2 s}$$

O laço L_1 não toca o laço L_2. (O laço L_1 toca o laço L_3 e o laço L_2 toca o laço L_3.) Portanto, o determinante Δ é dado por

$$\Delta = 1 - (L_1 + L_2 + L_3) + (L_1 L_2)$$
$$= 1 + \frac{1}{R_1 C_1 s} + \frac{1}{R_2 C_2 s} + \frac{1}{R_1 C_2 s} + \frac{1}{R_1 C_1 R_2 C_2 s^2}$$

Desde que todos os três laços tocam o caminho direto P_1, removemos L_1, L_2 e L_3 de Δ e calculamos o cofator Δ_1 como segue:

$$\Delta_1 = 1$$

Portanto obtemos a função de transferência em malha-fechada como segue:

$$\frac{H(s)}{Q(s)} = \frac{P_1 \Delta_1}{\Delta}$$

$$= \frac{\dfrac{1}{R_1 C_1 C_2 s^2}}{1 + \dfrac{1}{R_1 C_1 s} + \dfrac{1}{R_2 C_2 s} + \dfrac{1}{R_1 C_2 s} + \dfrac{1}{R_1 C_1 R_2 C_2 s^2}}$$

$$= \frac{R_2}{R_1 C_1 R_2 C_2 s^2 + (R_1 C_1 + R_2 C_2 + R_2 C_1)s + 1}$$

PROBLEMAS

Problema B.4.1 Obtenha as funções de transferência dos sistemas mecânicos indicados nas Figs. 4.55(a) e (b). Obtenha também os análogos elétricos dos sistemas mecânicos.

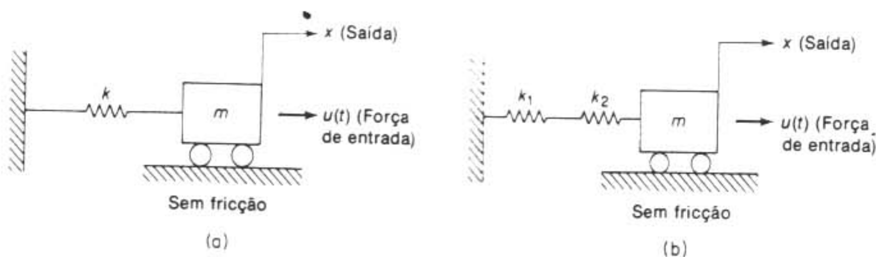


Fig. 4.55 Sistemas mecânicos.

Problema B.4.2 Obtenha a função de transferência de cada um dos três sistemas mecânicos indicados na Fig. 4.56. Nos diagramas, x_i denota o deslocamento de entrada e x_0 denota o deslocamento de saída. Suponha que os sistemas sejam a parâmetros concentrados (sem massa distribuída) e que o efeito de carga na saída é desprezível.

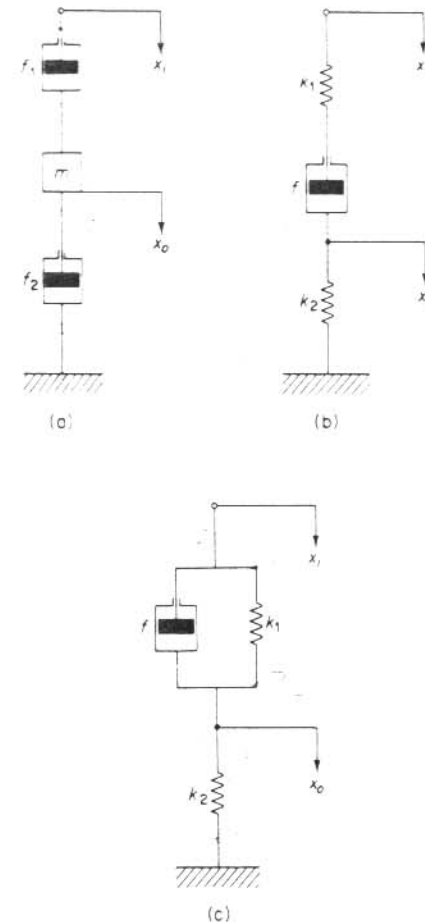
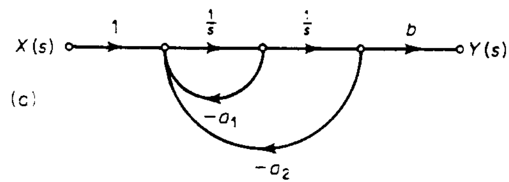
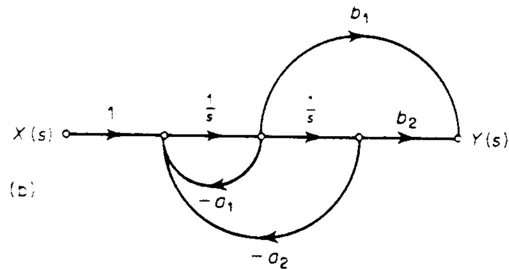


Fig. 4.56 Sistemas mecânicos.

Problema B.4.3 Obtenha as funções de transferência dos sistemas indicados nas Figs. 4.57(a), (b) e (c).





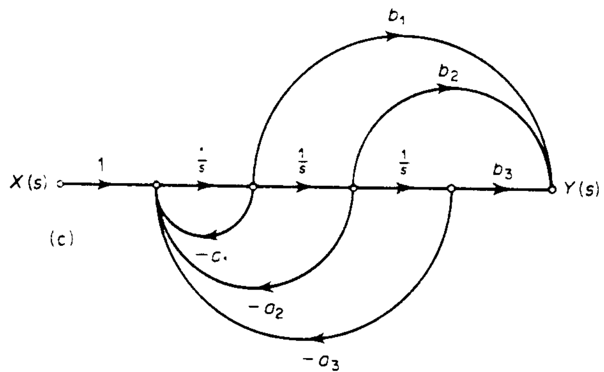
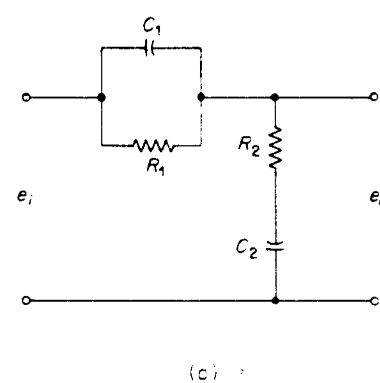


Fig. 4.57 Gráfico de fluxo de sinal de sistemas.

Problema B.4.4 Mostre que os sistemas indicados nas Figs. 4.58(a) e (b) são sistemas análogos. (Mostre que as funções de transferências dos dois sistemas possuem a mesma forma.)



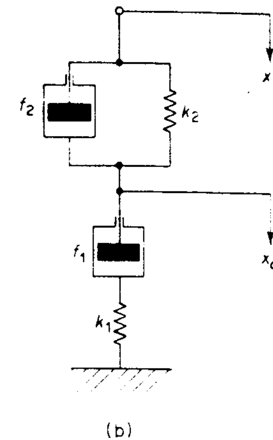


Fig. 4.58 (a) Sistema elétrico: (b) análogo mecânico do sistema elétrico.

Problema B.4.5 Deduza a função de transferência da rede elétrica mostrada na Fig. 4.59. Desenhe um diagrama esquemático de uma rede mecânica equivalente.

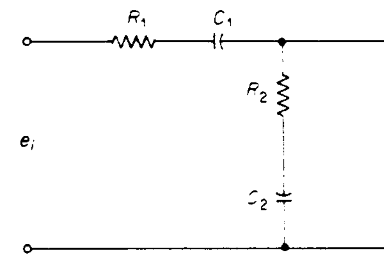


Fig. 4.59 Rede elétrica.

Problema B.4.6 Deduza a função de transferência da rede mecânica indicada na Fig. 4.60. Desenhe um diagrama esquemático de uma rede elétrica equivalente.

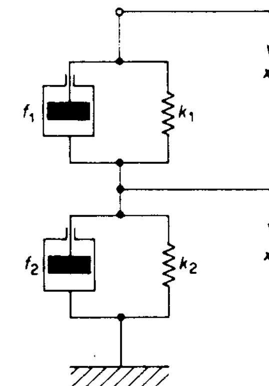


Fig. 4.60 Rede mecânica.

Problema B.4.7 Obtenha a função de transferência $\Theta(s)/E_f(s)$ do motor c.c. controlado por campo indicado na Fig. 4.61. No sistema, suponha que $J = 0.5$ lb-ft-s², $f = 0.2$ lb-ft/rad/s, e K_2 = constante de torque do motor = 27.4 lb-ft/amp.

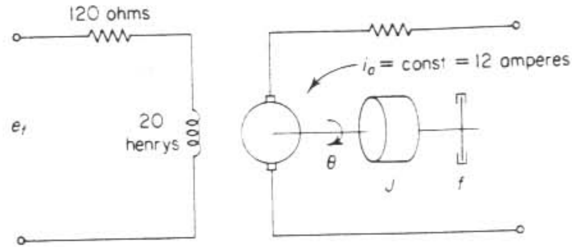


Fig. 4.61 Motor c.c. controlado por campo.

Problema B.4.8 Desenhe um diagrama de blocos do sistema de controle indicado na Fig. 4.62. Suponha que as funções de transferência do detector de erro *synchro* e do amplificador sejam constantes (K_1 e K_2, respectivamente) e a função de transferência do motor mais carga seja

$$\frac{C(s)}{E_a(s)} = \frac{K_3}{s(Ts + 1)}$$

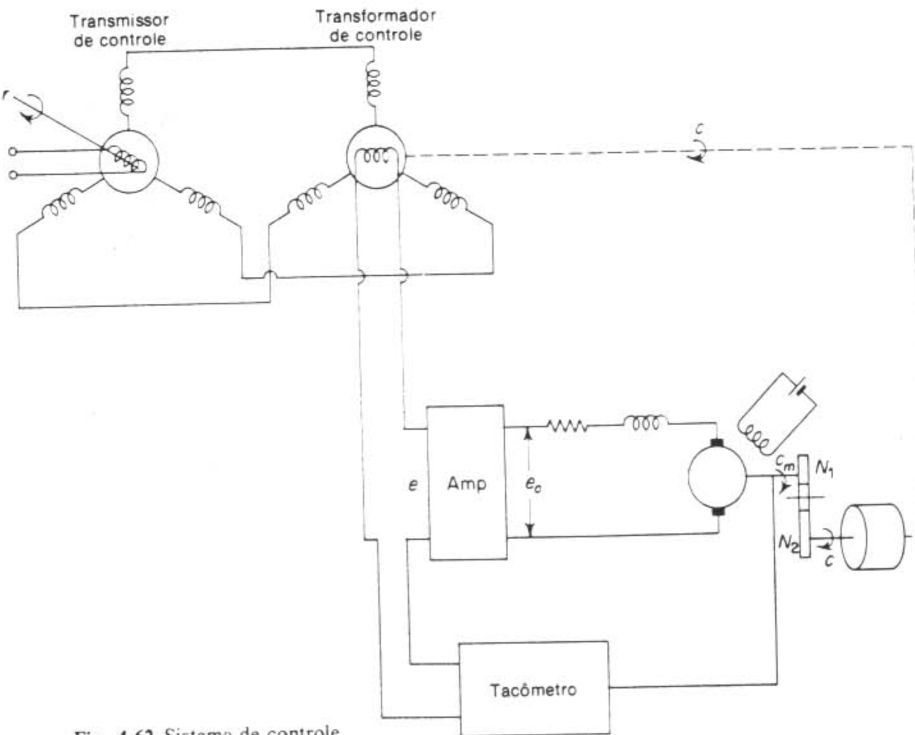


Fig. 4.62 Sistema de controle.

O coeficiente de realimentação de velocidade é K_4. A relação de engrenagem é $n = N_1/N_2 < 1$.

Problema B.4.9 Considere o sistema de nível de líquido indicado na Fig. 4.27(a). Supondo que $\bar{H} = 10$ ft, $\bar{Q} = 1$ ft³/s e que a área da seção transversal do tanque é igual a 5 ft², obtenha a constante de tempo do sistema no ponto de operação $(\bar{H}, \bar{Q})$.

Problema B.4.10 Obtenha a capacitância do recipiente de pressão que contém 2.500 in³ de ar a 100°F e 34.7 psia. Suponha que o processo de expansão é isotérmico.

Problema B.4.11 Obtenha a função de transferência do sistema de pressão de ar indicado na Fig. 4.63(a). Suponha que p_1 é a entrada e p_2 é a saída. O volume do recipiente é 5 in³ e a curva Δp *versus* fluxo é fornecida na Fig. 4.63(b). A temperatura do ar é 75°F e supõe-se que o processo de expansão do ar é isotérmico. Suponha que a pressão de operação $\Delta p = p_1 - p_2$ é 0 ~ 10 lb/in². (Use uma resistência média.)

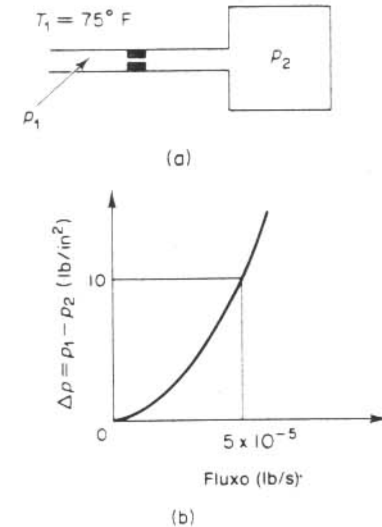


Fig. 4.63 (a) Sistema de pressão; (b) curva Δp *versus* fluxo.

Problema B.4.12 Obtenha a função de transferência do termômetro indicado na Fig. 4.64. Obtenha um análogo elétrico do termômetro.

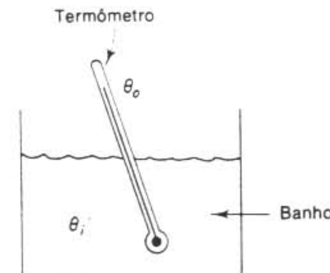


Fig. 4.64 Sistema termométrico.

Problema B.4.13 Um termopar possui uma constante de tempo de 2 segundos. Um recipiente térmico possui uma constante de tempo de 30 segundos. Quando o termopar é inserido no recipiente, este dispositivo de medida de temperatura pode ser considerado um sistema com dupla capacidade.

Determine as constantes de tempo do sistema conjunto termopar-recipiente térmico. Suponha que o peso do termopar é 0,5 oz e o peso do recipiente térmico é 3,2 oz. Suponha também que os calores específicos do termopar e do recipiente térmico sejam os mesmos.

Problema B.4.14 A Fig. 4.65 representa um diagrama de blocos de um sistema de controle de velocidade de uma máquina. A velocidade é medida por um conjunto de volantes. Desenhe um diagrama de fluxo de sinal para este sistema.

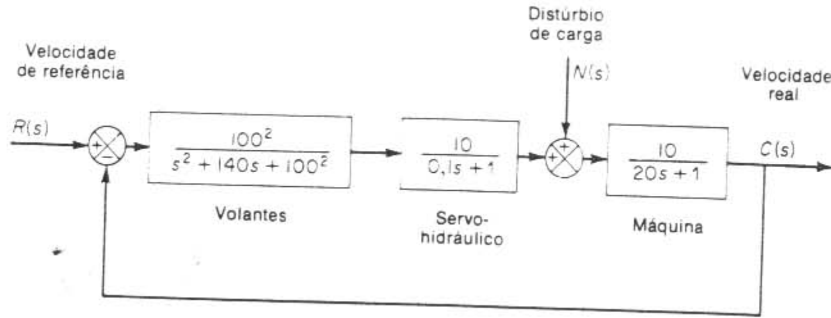


Fig. 4.65 Diagrama de blocos de um sistema de controle de velocidade de máquina.

5

Ações de Controle Básicas e Controles Automáticos Industriais

5.1 INTRODUÇÃO

Um controlador automático compara o valor real da saída do processo com o valor desejado, determina o desvio, e produz um sinal de controle que reduz o desvio a um valor nulo ou muito pequeno. A maneira pela qual o controlador automático produz o sinal de controle é denominada *ação de controle*.

Neste capítulo, apresentaremos as ações de controle básicas comumente utilizadas em controladores automáticos industriais. Inicialmente introduziremos o princípio de operação de controladores automáticos e os métodos para geração de vários sinais de controle, tais como o uso da derivada e da integral do sinal erro. Posteriormente discutiremos os efeitos de modos de controle particulares no desempenho do sistema. Então, daremos uma breve discussão de métodos para reduzir os efeitos de distúrbios externos no desempenho do sistema. Finalmente, introduziremos amplificadores fluídicos, apresentaremos princípios básicos da teoria de sistemas com fluidos e discutiremos aplicações de dispositivos que operam com fluidos.

Classificação de controladores automáticos industriais. Os controladores automáticos industriais podem ser classificados de acordo com a ação de controle como

1. controladores de duas posições ou liga-desliga (*on-off*)
2. controladores proporcionais
3. controladores do tipo integral
4. controladores do tipo proporcional-mais-integral
5. controladores do tipo proporcional-mais-derivativo
6. controladores do tipo proporcional-mais-derivativo-mais-integral

A maioria dos controladores automáticos industriais utilizam a eletricidade ou fluido pressurizado, tal como óleo ou ar, como fontes de potência. Os controladores automáticos também podem ser classificados de acordo com o tipo de potência empregado na operação, tais como controladores pneumáticos, controladores hidráulicos ou controladores eletrônicos. Qual dos tipos de controlador utilizar deve ser decidido pela natureza do processo e condições de operação, incluindo considerações como segurança, custo, disponibilidade, precisão, confiabilidade, peso e dimensão.

Elementos de controladores automáticos industriais. Um controlador automático deve detectar o sinal erro atuante, que é normalmente de baixo nível de potência, e amplificá-lo a um nível suficientemente alto. Portanto, é necessário um amplificador. A saída do controlador automático é ligada a um dispositivo de potência, tal como um motor ou válvula pneumática, um motor hidráulico ou um motor elétrico.

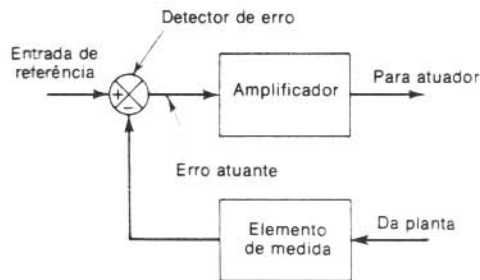


Fig. 5.1 Diagrama de blocos de um controlador automático industrial e elemento de medida.

A Fig. 5.1 mostra um diagrama de blocos de um controlador automático industrial conjuntamente com um elemento de medida. O controlador consiste em um detector de erro e amplificador. O elemento de medida é um dispositivo que converte a variável de saída em uma outra variável conveniente, tal como um deslocamento, pressão ou sinal elétrico que pode ser usada para comparar a saída em relação ao sinal de entrada de referência. Este elemento é o ramo de realimentação do sistema em malha-fechada. O ponto de ajuste do controlador deve ser convertido em uma entrada de referência com as mesmas unidades do sinal de realimentação na saída do elemento de medida. O amplificador amplifica a potência do sinal erro atuante, o qual por sua vez opera o atuador. (Na maioria das vezes, são usados um amplificador junto com um circuito de realimentação conveniente para operar o sinal erro atuante por amplificação e algumas vezes por diferenciação e/ou

integração a fim de produzir um melhor sinal de controle.) O atuador é um elemento que altera a entrada para o processo de acordo com o sinal de controle, de modo que o sinal realimentado possa ser comparado com o sinal de entrada de referência.

Controladores auto-operados. Em muitos controladores automáticos industriais são utilizadas unidades separadas para o elemento de medida e para o atuador. Em uma forma mais simples, entretanto, tal como em um controlador auto-operado, estes elementos são montados em uma única unidade. Controladores auto-operados utilizam a potência desenvolvida pelo elemento de medida e são muito simples e baratos. Um exemplo de um controlador auto-operado é mostrado na Fig. 5.2. O ponto de ajuste é determinado pelo ajuste da força na mola. A pressão controlada é medida pelo diafragma. O sinal erro atuante é a força atuante líquida no diafragma. Sua posição determina a abertura da válvula.

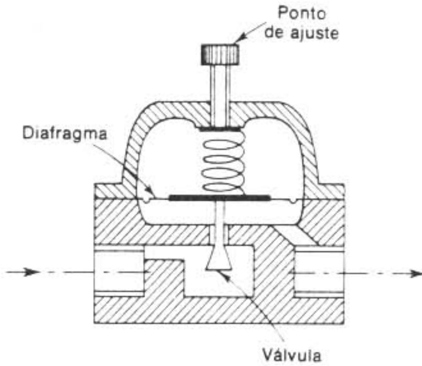


Fig. 5.2 Controlador auto-operado.

A operação do controlador auto-operado é a seguinte: Suponha que a pressão de saída é menor do que a pressão de referência, determinada pelo ponto de ajuste. Então, a força da mola para baixo é maior do que a força de pressão para cima, resultando em um movimento para baixo do diafragma. Isto acarreta um aumento da taxa de fluxo e aumento da pressão de saída. Quando a força de pressão para cima é igual à força da mola para baixo, a válvula permanece estacionária e a taxa de fluxo é constante. Inversamente, se a pressão de saída é maior do que a pressão de referência, a abertura da válvula torna-se menor e reduz a taxa de fluxo através da abertura da válvula. Este controlador auto-operado é muito usado para controle de pressão de água e gás. Neste controlador, a taxa de fluxo através da abertura da válvula é aproximadamente proporcional ao sinal erro atuante.

Ações de controle. As seis ações de controle básicas seguintes são muito comuns entre os controladores automáticos industriais: duas posições ou *on-off*, proporcional, integral, proporcional-mais-integral, proporcional-mais-derivativa e proporcional-mais-derivativa-mais-integral. Estas seis ações de controle serão discutidas neste capítulo. Note que é necessário um entendimento das características básicas das várias ações para que o engenheiro de controle possa selecionar a mais conveniente em sua particular aplicação.

Ação de controle de duas posições ou liga-desliga *(on-off)*. Em um sistema de controle de duas posições, o elemento atuante possui apenas duas posições fixas que são, em muitos casos, simplesmente ligado e desligado (do inglês, *on-off*). O controle de duas posições é relativamente simples e barato e, por esta razão, extremamente utilizado tanto em sistemas de controle industriais como domésticos.

Considere o sinal de saída do controlador $m(t)$ e o sinal erro atuante $e(t)$. Em um controle de duas posições, o sinal $m(t)$ permanece ou em um valor máximo ou em um valor mínimo, dependendo de o sinal erro atuante ser positivo ou negativo, de modo que

$$m(t) = M_1 \quad \text{para } e(t) > 0$$
$$= M_2 \quad \text{para } e(t) < 0$$

onde M_1 e M_2 são constantes. O valor mínimo M_2 usualmente ou é zero ou $-M_1$. Os controladores de duas posições geralmente são dispositivos elétricos, e uma válvula operada por solenóide elétrico é extensivamente usada nestes controladores. Controladores proporcionais pneumáticos com ganhos muito altos atuam como controladores de duas posições e são muitas vezes denominados controladores de duas posições pneumáticos.

As Figs. 5.3(a) e (b) mostram os diagramas de blocos para controladores de duas posições. O intervalo através do qual o sinal erro atuante deve mover-se antes de ocorrer o chaveamento é denominado intervalo diferencial. Um intervalo *(gap)* diferencial é indicado na Fig. 5.3(b). Este intervalo diferencial faz com que a saída

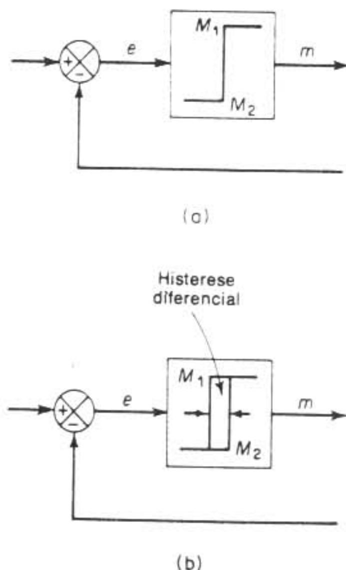


Fig. 5.3 (a) Diagrama de blocos de um controlador liga-desliga; (b) diagrama de blocos de um controlador liga-desliga com histerese diferencial.

do controlador $m(t)$ mantenha seu valor presente até que o sinal erro atuante tenha-se movido ligeiramente além do valor zero. Em alguns casos, o intervalo diferencial é um resultado de fricção e perda de movimento não intencionalmente introduzido no sistema; entretanto, normalmente é intencionalmente colocado de modo a prevenir uma operação freqüente demais do mecanismo *on-off*.

Considere o sistema de controle de nível de líquido indicado na Fig. 5.4. Com o controle de duas posições, a válvula ou está aberta ou está fechada. Conseqüentemente a taxa de fluxo de entrada de água ou é uma constante positiva ou é zero. Conforme indicado na Fig. 5.5, o sinal de saída move-se continuamente entre os dois limites exigidos de modo a fazer com que o elemento atuante mova-se de uma posição fixa para outra. Note que a curva de saída segue uma de duas curvas exponenciais, uma correspondendo à curva para encher e a outra correspondendo à curva para esvaziar. Esta oscilação da saída entre dois limites é uma característica de resposta típica de um sistema sob controle de duas posições.

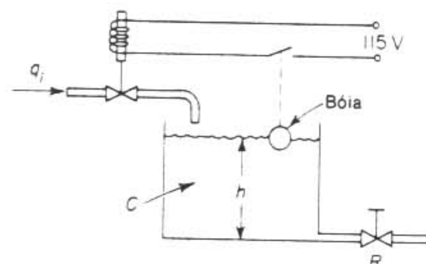


Fig. 5.4 Sistema de controle de nível de líquido.

Da Fig. 5.5, verificamos que a amplitude da oscilação de saída pode ser reduzida diminuindo-se o intervalo diferencial. Isto, entretanto, aumenta o número de operações da chave, ligando-a e desligando-a por minuto, e reduz a vida útil do componente. O tamanho do intervalo diferencial deve ser determinado a partir de considerações relativas à precisão exigida e à vida do componente. Análises posteriores de sistemas de controle com controladores de duas posições são deixadas para o Cap. 11.

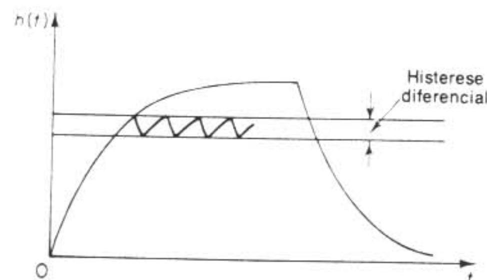


Fig. 5.5 Curva de nível $h(t)$ *versus* t para o sistema indicado na Fig. 5.4.

Ação de controle proporcional. Para um controlador com ação de controle proporcional, a relação entre a saída do controlador $m(t)$ e o sinal erro atuante $e(t)$ é

$$m(t) = K_p e(t)$$

ou, em transformada de Laplace,

$$\frac{M(s)}{E(s)} = K_p$$

onde K_p é denominado sensibilidade proporcional ou ganho.

Independentemente do mecanismo real, ou da forma da potência de operação, o controlador proporcional é essencialmente um amplificador com um ganho ajustável. Um diagrama de blocos deste controlador é mostrado na Fig. 5.6.

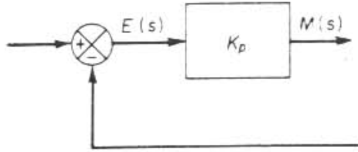


Fig. 5.6 Diagrama de blocos de um controlador proporcional.

Ação de controle integral. Em um controlador com a ação de controle integral, o valor da saída do controlador $m(t)$ é variado em uma taxa proporcional ao sinal erro atuante $e(t)$. Isto é,

$$\frac{dm(t)}{dt} = K_i e(t)$$

ou

$$m(t) = K_i \int_0^t e(t)\,dt$$

onde K_i é uma constante ajustável. A função de transferência do controlador integral é

$$\frac{M(s)}{E(s)} = \frac{K_i}{s}$$

Se o valor de $e(t)$ é dobrado, então o valor de $m(t)$ varia duas vezes mais rápido. Para erro atuante nulo, o valor de $m(t)$ permanece estacionário. A ação de controle integral é muitas vezes denominada controle de restabelecimento (*reset*). A Fig. 5.7 mostra um diagrama de blocos deste tipo de controlador.

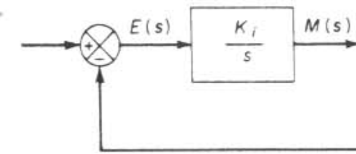


Fig. 5.7 Diagrama de blocos de um controlador integral.

Ação de controle proporcional-mais-integral. A ação de controle de um controlador proporcional-mais-integral é definida pela seguinte equação:

$$m(t) = K_p e(t) + \frac{K_p}{T_i}\int_0^t e(t)\,dt$$

ou a função de transferência do controlador é

$$\frac{M(s)}{E(s)} = K_p\left(1 + \frac{1}{T_i s}\right)$$

onde K_p representa a sensibilidade proporcional ou ganho e T_i representa o tempo integral. Tanto K_p como T_i são ajustáveis. O tempo integral ajusta a ação de controle integral, enquanto uma mudança no valor de K_p afeta tanto a parte proporcional como a parte integral da ação de controle. O inverso do tempo integral

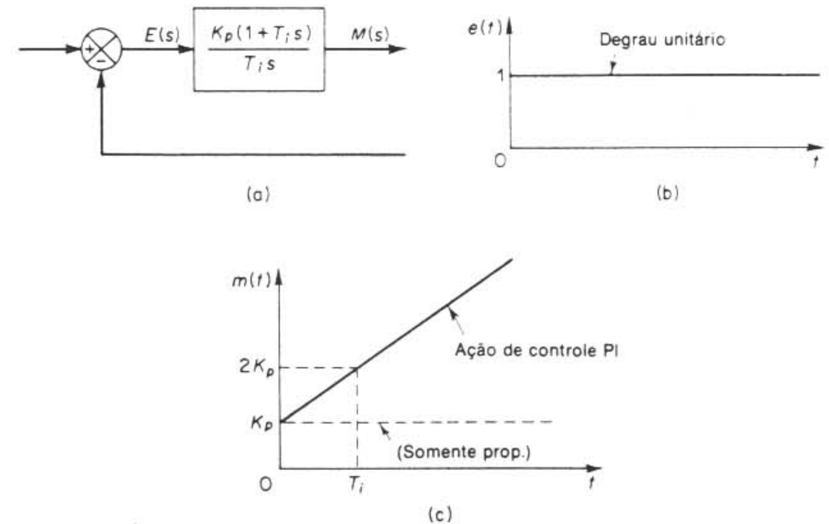
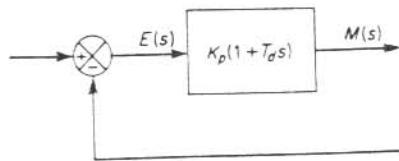
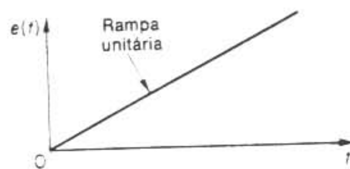


Fig. 5.8 (a) Diagrama de blocos de um controlador proporcional-mais-integral; (b) e (c) diagramas indicando a entrada com degrau unitário e a saída do controlador.





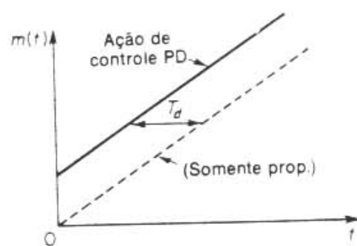


Fig. 5.9 (a) Diagrama de blocos de um controlador proporcional-mais-derivativo; (b) e (c) diagramas indicando a entrada em rampa unitária e a saída do controlador.

T_i é denominado taxa de restabelecimento (*reset*). A taxa de restabelecimento é o número de vezes por minuto que a parte proporcional da ação de controle é duplicada. A taxa de restabelecimento é medida em termos de repetições por minuto. A Fig. 5.8(a) mostra um diagrama de blocos de um controlador proporcional-mais-integral. Se o sinal erro atuante $e(t)$ é uma função degrau unitário, como mostrado na Fig. 5.8(b), então a saída do controlador $m(t)$ é a indicada na Fig. 5.8(c).

Ação de controle proporcional-mais-derivativa. A ação de controle de um controlador proporcional-mais-derivativo é definida pela seguinte equação:

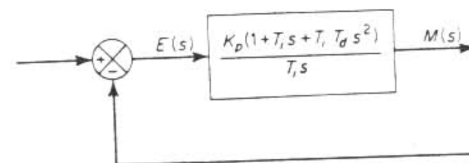
$$m(t) = K_p e(t) + K_p T_d \frac{de(t)}{dt}$$

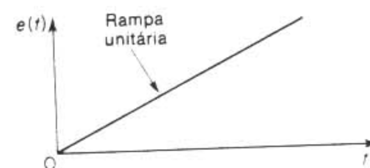
e a função de transferência é

$$\frac{M(s)}{E(s)} = K_p(1 + T_d s)$$

onde K_p representa a sensibilidade proporcional e T_d representa o tempo derivativo. Tanto K_p como T_d são ajustáveis. A ação de controle derivativa, algumas vezes denominada controle de taxa, é onde a magnitude da saída do controlador é proporcional à taxa de variação do sinal erro atuante. O tempo derivativo T_d é o intervalo de tempo pelo qual a ação de taxa avança o efeito da ação de controle proporcional. A Fig. 5.9(a) mostra um diagrama de blocos de um controlador proporcional-mais-derivativo. Se o sinal erro atuante $e(t)$ é uma função rampa unitária conforme mostrado na Fig. 5.9(b), então a saída do controlador $m(t)$ é a indicada na Fig. 5.9(c). Como pode ser visto da Fig. 5.9(c), a ação de controle derivativa tem um caráter antecipatório. De fato, entretanto, a ação de controle derivativa nunca pode antecipar uma ação que ainda não ocorreu.

Enquanto a ação de controle derivativa possui a vantagem de ser antecipatória, tem as desvantagens de amplificar os sinais de ruído e causar um efeito de saturação no atuador.





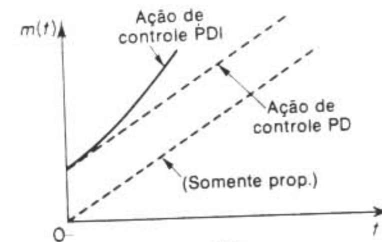


Fig. 5.10 (a) Diagrama de blocos de um controlador proporcional-mais-derivativo-mais-integral; (b) e (c) diagramas indicando a entrada em rampa unitária e a saída do controlador.

Note que a ação de controle derivativa nunca pode ser usada sozinha porque esta ação somente é efetiva durante os intervalos de tempo correspondentes a transitórios.

Ação de controle proporcional-mais-derivativa-mais-integral. A combinação da ação de controle proporcional, ação de controle derivativa e ação de controle integral é denominada ação de controle proporcional-mais-derivativa-mais-integral. Esta ação combinada possui as vantagens de cada uma das três ações de controle individuais. A equação de um controlador com esta ação combinada é dada por

$$m(t) = K_p e(t) + K_p T_d \frac{de(t)}{dt} + \frac{K_p}{T_i}\int_0^t e(t)\,dt$$

ou a função de transferência é

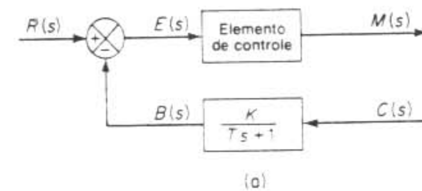
$$\frac{M(s)}{E(s)} = K_p\left(1 + T_d s + \frac{1}{T_i s}\right)$$

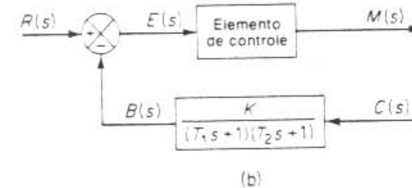
onde K_p representa a sensibilidade proporcional, T_d representa o tempo derivativo e T_i representa o tempo integral. O diagrama de blocos de um controlador proporcional-mais-derivativo-mais-integral é mostrado na Fig. 5.10(a). Se $e(t)$ é uma função rampa unitária conforme mostrado na Fig. 5.10(b), então a saída do controlador $m(t)$ é a indicada na Fig. 5.10(c).

Efeitos do elemento de medida no desempenho do sistema. Desde que as características estática e dinâmica do elemento de medida afetam a indicação do valor real da variável de saída, o elemento de medida representa um papel importante na determinação do desempenho global do sistema de controle. O elemento de medida normalmente determina a função de transferência no ramo de realimentação. Se as constantes de tempo do elemento de medida são suficientemente pequenas, comparadas com outras constantes de tempo no sistema de controle, a função de transferência do elemento de medida simplesmente torna-se uma constante. As Figs. 5.11(a), (b) e (c) mostram diagramas de blocos de controladores automáticos contendo um elemento de medida de primeira-ordem, de segunda-ordem sobreamortecido e segunda-ordem subamortecido, respectivamente. A resposta de um elemento de medida térmico é normalmente do tipo de segunda-ordem sobreamortecido.

Diagramas de blocos de sistemas de controle automático. Um diagrama de bloco de um sistema de controle automático pode ser obtido pela ligação da planta ao controlador automático, como mostrado na Fig. 5.12. A realimentação do sinal de saída é obtida pelo elemento de medida. A equação relacionando a variável de saída $C(s)$ com a entrada de referência $R(s)$ e a variável de distúrbio $N(s)$ pode ser obtida como segue:

$$C(s) = \frac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)H(s)}R(s) + \frac{G_3(s)}{1 + G_1(s)G_2(s)H(s)}N(s)$$





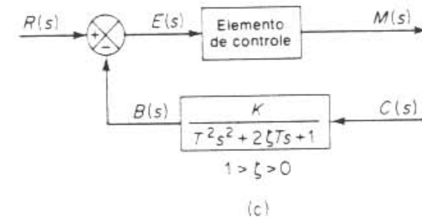


Fig. 5.11 Diagramas de blocos de controladores automáticos com (a) elemento de medida de primeira-ordem; (b) elemento de medida de segunda-ordem superamortecido; (c) elemento de medida de segunda-ordem subamortecido.

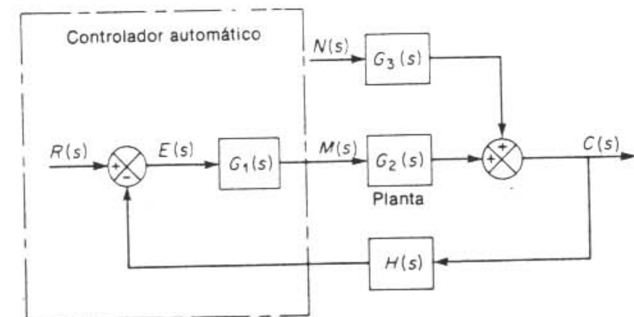


Fig. 5.12 Diagrama de blocos de um sistema de controle.

Nos sistemas de controle de processos, usualmente estamos interessados na resposta ao distúrbio $N(s)$. Em servomecanismo, entretanto, a resposta a uma entrada de referência $R(s)$ é de maior interesse. Iremos apresentar a análise da resposta do sistema em relação a variações nos distúrbios e na carga na Seção 5.4. A resposta do sistema em relação a variações na entrada de referência será estudada em detalhes no Cap. 6.

5.2 CONTROLADORES PROPORCIONAIS

Nesta seção, ilustraremos o fato de que controladores proporcionais utilizam o princípio da realimentação negativa neles próprios. Forneceremos uma discussão detalhada do princípio pelo qual operam os controladores proporcionais, considerando um controlador pneumático. Mostraremos então que o mesmo princípio aplica-se a controladores hidráulicos e eletrônicos. Através desta discussão, daremos maior ênfase aos princípios fundamentais do que a detalhes da operação de mecanismos reais.

Sistemas pneumáticos. Controladores pneumáticos de baixa pressão têm sido desenvolvidos para sistemas de controle industriais e extensivamente usados em processos industriais. As razões para o amplo uso de controladores pneumáticos são, principalmente, suas características a prova de explosão, simplicidade e facilidade de manutenção.

Amplificadores pneumáticos bocal-haste *(nozzle-flapper)*. Um diagrama esquemático de um amplificador pneumático bocal-haste (*nozzle-flapper*) é indicado na Fig. 5.13(a). A fonte de potência para este amplificador é uma fonte de ar a pressão constante. O amplificador converte pequenas variações na posição da haste em grandes variações na pressão do bocal. Portanto uma grande potência de saída pode ser controlada através de uma pequena potência necessária para posicionar a haste (*flapper*).

Na Fig. 5.13(a) o ar pressurizado é introduzido através de um orifício, e o ar é ejetado do bocal (*nozzle*) para a haste. Usualmente a pressão de alimentação P_s para este controlador é de 20 psig. O diâmetro do orifício é da ordem de 0.25 mm e o do bocal é da ordem de 0.4 mm. O diâmetro do bocal deve ser maior que o do orifício, a fim de que o amplificador funcione apropriadamente. A haste é posicionada contra a abertura do bocal e a pressão no bocal P_b é controlada pela distância X haste-bocal. Conforme a haste se aproxima do bocal, a oposição para o fluxo de ar através do bocal aumenta, resultando em um aumento da pressão P_b. Se o bocal for completamente fechado pela haste, a pressão P_b do bocal é igual à pressão de suprimento P_s. Se a haste for distanciada do bocal, de modo que a distância bocal-haste seja grande (da ordem de 0.25 mm), então não há praticamente restrição ao fluxo e a pressão P_b no bocal possui o seu valor mínimo que depende do dispositivo haste-bocal utilizado. (A mínima pressão possível será a pressão ambiente P_a.)

Notar que, como o jato de ar empurra a haste, é necessário construir o diâmetro do bocal tão pequeno quanto possível.

Uma curva típica relacionando a pressão P_b do bocal com a distância haste-bocal X é mostrada na Fig. 5.13(b). A parte descendente, quase linear, da curva é utilizada na operação real do amplificador haste-bocal. Como o intervalo de deslocamentos da haste é restrito a valores muito pequenos, a variação na pressão de saída também é pequena, apesar de a curva aparentar ser muito íngreme.

O amplificador haste-bocal converte deslocamento em um sinal de pressão. Desde que os sistemas de controle de processo industrial exigem grandes potências de saída para operarem grandes válvulas atuantes pneumáticas, usualmente a

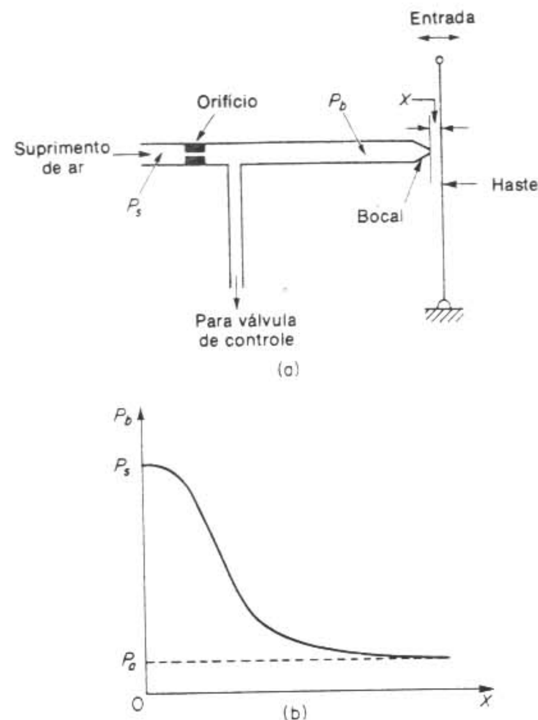


Fig. 5.13 (a) Diagrama esquemático de um amplificador pneumático bocal-haste; (b) curva característica relacionando a pressão no bocal e a distância bocal-haste.

amplificação de potência do amplificador haste-bocal não é suficiente. Conseqüentemente, quase sempre é empregado um relé pneumático como amplificador de potência em conexão com o amplificador haste-bocal.

Relés pneumáticos. Em um controlador pneumático prático, um amplificador haste-bocal atua como primeiro-estágio amplificador e um relé pneumático como o segundo-estágio amplificador. O relé pneumático é capaz de operar com uma grande quantidade de fluxo de ar.

Um diagrama esquemático de um relé pneumático é indicado na Fig. 5.14(a). Conforme a pressão no bocal P_b aumenta, a válvula de bola é forçada para a posição inferior, diminuindo a pressão de controle P_c. Um relé deste tipo é denominado relé de ação reversa.

Quando a válvula de bola está na posição superior, a abertura atmosférica é fechada e a pressão de controle P_c torna-se igual à pressão de suprimento P_s. Quando a válvula de bola está na parte mais inferior de sua posição, ela fecha completamente o suprimento de ar e a pressão de controle P_c cai para a pressão atmosférica. A pressão de controle P_c pode então ser variada desde o número 0 psig até a pressão total de suprimento, usualmente 20 psig.

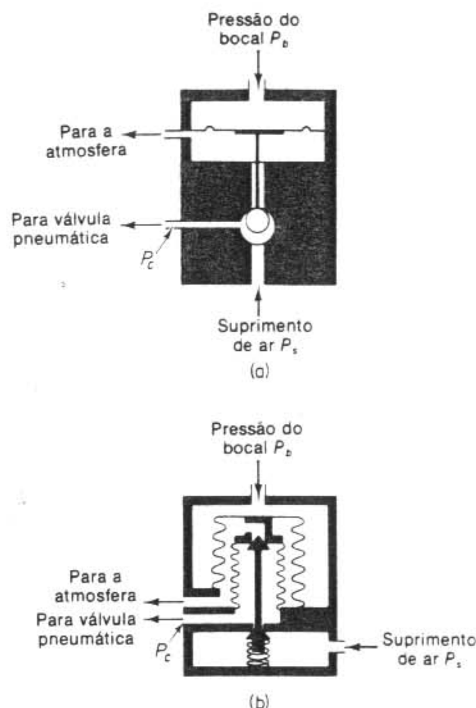


Fig. 5.14 (a) Diagrama esquemático de um relé tipo-sangria; (b) diagrama esquemático de um relé tipo sem sangria.

O movimento total da válvula de bola entre as posições mais inferior e mais superior é muito pequeno (da ordem de 0,25 mm). Em todas as posições da válvula de bola, exceto na posição mais superior, o ar continua a escapar para a atmosfera, mesmo após ser atingida a condição de equilíbrio entre a pressão do bocal e a pressão de controle. Por este motivo o relé mostrado na Fig. 5.14(a) é denominado um relé tipo sangria ("escape").

Há um outro tipo de relé, o tipo sem sangria. Neste tipo o ar deixa de escapar quando a condição de equilíbrio é atingida e, portanto, não há perda de ar pressurizado na operação em estado estacionário. Note, entretanto, que o relé tipo sem sangria deve possuir uma saliência para libertar a pressão de controle P_c da válvula atuante pneumática. Um diagrama esquemático de um relé tipo sem sangria é mostrado na Fig. 5.14(b).

Em qualquer tipo de relé o suprimento de ar é controlado por uma válvula, que por sua vez é controlada pela pressão do bocal. Portanto, a pressão do bocal é convertida na pressão de controle com amplificação de potência.

Desde que a pressão de controle P_c varia quase instantaneamente com variações na pressão do bocal P_b, a constante de tempo do relé pneumático é desprezível comparada com as outras constantes de tempo maiores do controlador pneumático e da planta.

Controladores proporcionais pneumáticos (tipo força-distância). Dois tipos de controladores pneumáticos, um denominado do tipo força-distância e outro do tipo força-balanço, são extensivamente utilizados na indústria. Indiferentemente de quão distintos os controladores pneumáticos industriais possam parecer, um estudo cuidadoso mostrará a grande similaridade nas funções do circuito pneumático. Neste livro consideraremos apenas os controladores pneumáticos tipo força-distância.

A Fig. 5.15(a) mostra um diagrama esquemático deste tipo de controlador proporcional. O amplificador haste-bocal constitui o primeiro estágio amplificador, e a pressão do bocal é controlada pela distância haste-bocal. O amplificador tipo relé constitui o segundo estágio amplificador. A pressão do bocal determina a posição da válvula de mola para o segundo estágio amplificador, que é capaz de operar com uma grande quantidade de fluxo de ar.

Na maioria dos controladores pneumáticos, é empregado algum tipo de realimentação pneumática. A realimentação da saída pneumática reduz a quantidade de movimento real da haste. Ao invés de montar a haste em um ponto fixo, como indicado na Fig. 5.15(b), ela é normalmente pivotada no fole de realimentação, como indicado na Fig. 5.15(c). A quantidade de realimentação pode ser regulada pela introdução de uma ligação variável entre o fole de realimentação e o ponto de conexão da haste. A haste torna-se então um elo flutuante. Ela pode mover-se tanto por ação do sinal erro como pelo sinal de realimentação.

A operação do controlador indicado na Fig. 5.15(a) é a seguinte: O sinal de entrada para os dois estágios amplificadores pneumáticos é o sinal erro atuante. Um aumento no sinal erro atuante move a haste para a direita. Esta, por sua vez, diminui a pressão no bocal, e o fole B se contrairá, resultando em um movimento para cima da válvula de bola. Isto acarretará um fluxo maior para a válvula pneumática e a pressão de controle aumentará. Este aumento causa uma expansão do fole F e move a haste para a esquerda, fechando o bocal.

O deslocamento haste-bocal é muito pequeno devido a esta realimentação, mas a variação na pressão de controle pode ser grande. No caso em que o erro atuante decresce, a pressão no bocal aumenta e a válvula de bola move-se para baixo, resultando em um decréscimo no fluxo de suprimento para a válvula e um aumento na sangria para a atmosfera. Isto acarretará uma diminuição da pressão de controle.

É importante notar que o fole de realimentação moveria a haste de uma distância menor do que aquela causada apenas pelo sinal erro. Se estes dois movimentos fossem iguais, não resultaria uma ação de controle.

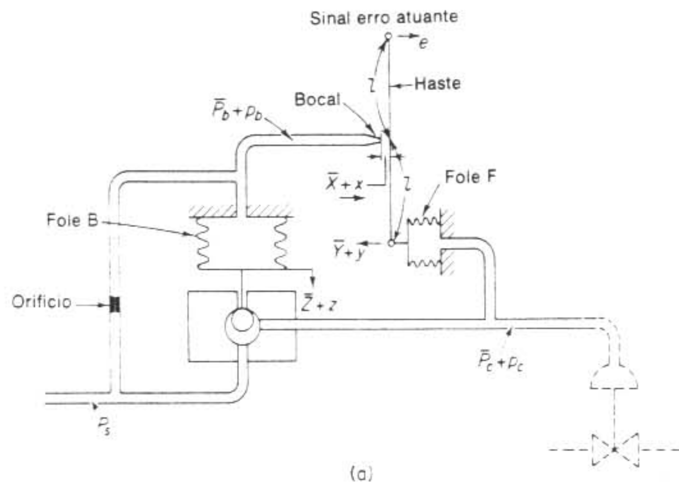
As equações para este controlador podem ser deduzidas como segue: Quando o erro atuante é nulo, ou $e = 0$, e existe um estado de equilíbrio com a distância bocal-haste igual a $\bar{X}$, o deslocamento do fole F é igual a $\bar{Y}$, o deslocamento do fole B igual a $\bar{Z}$, a pressão do bocal igual a $\bar{P}_b$, e a pressão de controle é igual a $\bar{P}_c$. Quando houver qualquer erro atuante, a distância haste-bocal, os deslocamentos dos foles F e B, a pressão no bocal e a pressão de controle desviam de seus respectivos valores de equilíbrio. Consideremos estes desvios como sendo x, y, z, p_b, e p_c, respectivamente. (O sentido positivo para cada variável de deslocamento é indicado por uma seta.)

Supondo que a relação entre a variação na pressão do bocal e a variação na distância haste-bocal seja linear, temos

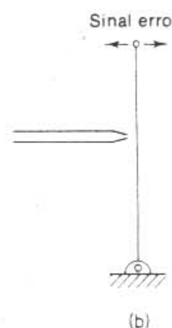
$$p_b = -K_1 x \tag{5-1}$$

onde K_1 é uma constante. Para o fole B,

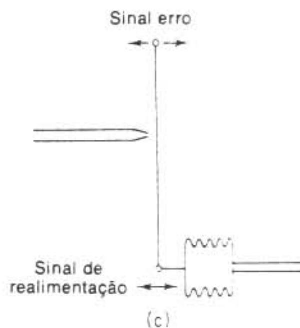
$$p_b = K_2 z \tag{5-2}$$



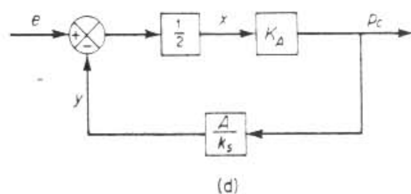
(a)



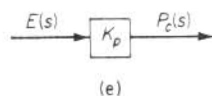
(b)



(c)



(d)



(e)

Fig. 5.15 (a) Diagrama esquemático de um controlador proporcional pneumático tipo força-distância; (b) haste montada em um ponto fixo; (c) haste montada no fole de realimentação; (d) diagrama de blocos do controlador; (e) diagrama de blocos simplificado para o controlador.

onde K_2 é uma constante. A posição da válvula de bola que depende do deslocamento do fole B determina a pressão de controle. Se a válvula de bola for tal que a relação entre p_c e z seja linear, então

$$p_c = -K_3 z \tag{5-3}$$

onde K_3 é uma constante. Das Eqs. (5.1), (5.2) e (5.3), obtemos

$$p_c = -\frac{K_3}{K_2} p_b = K_A x \tag{5-4}$$

onde $K_A = K_1 K_3 / K_2$ é uma constante. Para o movimento da haste, temos

$$x = \frac{e - y}{2} \tag{5-5}$$

O fole F age como uma mola, e vale a seguinte equação:

$$A p_c = k_s y \tag{5-6}$$

onde A é a área efetiva do fole F, e k_s é a constante da mola equivalente ou a rigidez devida à ação do lado corrugado do fole.

Supondo que todas as variações nas variáveis estão dentro da faixa linear, podemos obter um diagrama de blocos para este sistema a partir das Eqs. (5.4), (5.5) e (5.6), como indicado na Fig. 5.15(d). Da Fig. 5.15(d), pode ser verificado claramente que o controlador pneumático mostrado na Fig. 5.15(a) é por si só um sistema realimentado. A função de transferência entre p_c e e é dada por

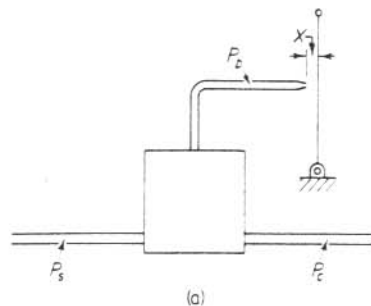
$$\frac{P_c(s)}{E(s)} = \frac{\frac{1}{2} K_A}{1 + \frac{K_A A}{2k_s}} = K_p \tag{5-7}$$

Um diagrama de blocos simplificado é mostrado na Fig. 5.15(e). Desde que p_c e e são proporcionais, o controlador pneumático indicado na Fig. 5.15(a) é denominado um controlador proporcional pneumático.

Note que desde que o valor de $K_A A / k_s$ é geralmente muito maior do que a unidade em controladores reais, a função de transferência dada pela Eq. (5.7) pode ser simplificada para resultar

$$\frac{P_c(s)}{E(s)} = \frac{\frac{1}{2} K_A}{\frac{K_A A}{2k_s}} = \frac{k_s}{A} \tag{5-8}$$

Como visto a partir das Eqs. (5.7) e (5.8), o ganho do controlador proporcional pneumático pode ser variado amplamente pelo ajuste da ligação conectando a haste. [A ligação conectando a haste não é indicada na Fig. 5.15(a).] Na maioria dos controladores proporcionais comerciais, existe um botão de ajuste, ou outro mecanismo para variar o ganho através do ajuste desta ligação.



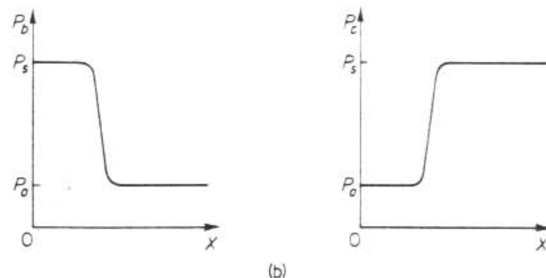


Fig. 5.16 (a) Controlador pneumático sem mecanismo de realimentação; (b) curvas P_b *versus* X e P_c *versus* X.

Como foi verificado anteriormente, o sinal erro atuante moveu a haste em um sentido e o fole de realimentação moveu a haste no sentido oposto, porém com um deslocamento menor. O efeito do fole de realimentação é, portanto, reduzir a sensibilidade do controlador. O princípio da realimentação é comumente usado para obter controladores de banda-proporcional ampla.

Os controladores pneumáticos que não possuem mecanismos de realimentação [o que significa que um extremo da haste é fixo, como indicado na Fig. 5.16(a)] possuem alta sensibilidade e são denominados controladores proporcionais de banda estreita ou controladores de duas posições. Neste tipo de controlador é exigido apenas um pequeno movimento entre o bocal e a haste a fim de obter-se uma variação completa desde a pressão de controle máxima até a mínima. As curvas relacionando P_b com X e P_c com X são mostradas na Fig. 5.16(b). Note que uma pequena variação em X pode acarretar uma grande variação em P_b, o que faz a válvula de bola estar completamente aberta ou completamente fechada.

Válvulas atuantes pneumáticas. Uma característica de controles pneumáticos é que eles empregam quase exclusivamente válvulas atuantes pneumáticas. Uma válvula atuante pneumática pode fornecer uma grande potência de saída. (Desde que um atuador pneumático requer uma grande potência na entrada para produzir uma grande potência na saída, é necessário que seja disponível uma quantidade suficiente de ar pressurizado.) Na prática, as válvulas atuantes pneumáticas possuem características que podem ser não lineares; isto é, o fluxo pode não ser

diretamente proporcional à posição da haste da válvula, e também pode haver outros efeitos não lineares, tal como histerese.

Considere o diagrama esquemático de uma válvula atuante pneumática indicado na Fig. 5.17. Suponha que a área do diafragma é A. Suponha ainda que, quando o erro atuante é nulo, a pressão de controle é igual a $\bar{P}_c$ e o deslocamento da válvula é igual a $\bar{X}$.

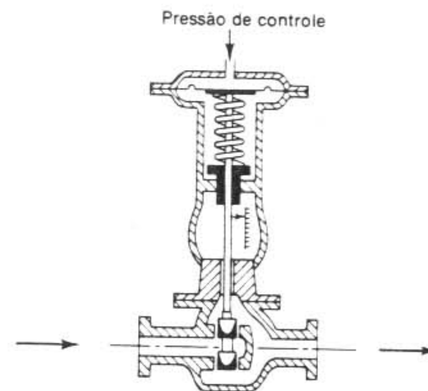


Fig. 5.17 Diagrama esquemático de uma válvula atuante pneumática.

Na análise seguinte, consideraremos pequenas variações nas variáveis e linearizaremos a válvula atuante pneumática. Vamos definir a pequena variação na pressão de controle e o correspondente deslocamento da válvula como sendo p_c e x, respectivamente. Desde que uma pequena variação na força de pressão pneumática aplicada ao diafragma reposiciona a carga, consistindo na mola, fricção-viscosa e massa, a equação de balanço de forças resulta

$$Ap_c = m\ddot{x} + f\dot{x} + kx \tag{5-9}$$

onde

m = massa da válvula e haste da válvula
f = coeficiente de fricção-viscosa
k = constante da mola

Se a força devida à massa e a fricção-viscosa são desprezíveis, então a Eq. (5.9) pode ser simplificada para

$$Ap_c = kx$$

A função de transferência entre x e p_c torna-se

$$\frac{X(s)}{P_c(s)} = \frac{A}{k} = K_c$$

onde $X(s) = \mathcal{L}[x]$ e $P_c(s) = \mathcal{L}[p_c]$. Se q_i, a variação no fluxo através da válvula atuante pneumática, for proporcional a x, a variação no deslocamento da haste da

válvula, então,

$$\frac{Q_i(s)}{X(s)} = K_q$$

onde $Q_i(s) = \mathscr{L}[q_i]$ e K_q é uma constante. A função de transferência entre q_i e p_c torna-se

$$\frac{Q_i(s)}{P_c(s)} = K_c K_q = K_v$$

onde K_v é uma constante.

A pressão de controle típica para este tipo de válvula atuante pneumática é de 3 a 15 psig. O deslocamento da haste da válvula é limitado pelo golpe possível no diafragma que é de poucos centímetros. Se um golpe maior for necessário, pode ser empregada uma combinação pistão-mola.

Em válvulas atuantes pneumáticas, a força de fricção estática deve ser limitada a um valor baixo de modo a não resultar em histerese excessiva. Devido à compressibilidade do ar, a ação de controle não pode ser positiva; isto é, deve existir um erro na posição da válvula de haste. O uso de um posicionador de válvula acarreta melhorias no desempenho de uma válvula atuante pneumática.

Um diagrama esquemático de um posicionador de válvula é indicado na Fig. 5.18. O princípio de operação deste dispositivo é que, se a posição da válvula não corresponder à pressão de controle, então a válvula piloto operará até que a posição da válvula corresponda exatamente à pressão de controle.

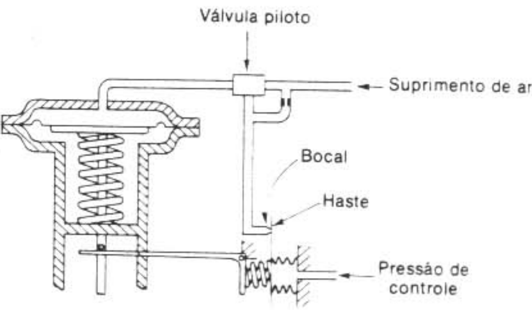


Fig. 5.18 Diagrama esquemático de um posicionador de válvula.

Sistemas de controle de nível de líquido. Considere o sistema de controle de nível de líquido indicado na Fig. 5.19. É desejado manter-se o nível de líquido em um valor constante, independentemente de variações na abertura da válvula de carga.

Suponhamos que o controlador seja do tipo proporcional, conforme mostrado na Fig. 5.15(a), e que a válvula de controle seja aquela mostrada na Fig. 5.17. Se o nível do líquido aumentar como resultado de uma variação na válvula de carga, a bóia move-se para cima, fazendo com que a haste mova-se mais próxima do contato com o bocal, aumentando a pressão do bocal. Desde que o relé é reversamente

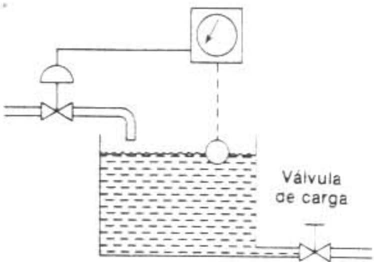


Fig. 5.19 Sistema de controle de nível de líquido.

atuante, resultará em uma redução na pressão de controle e causará um decréscimo na abertura da válvula atuante pneumática. Isto está no sentido apropriado para corrigir o nível crescente.

Controle proporcional de um sistema de primeira ordem. Considere o sistema de controle de nível de líquido indicado na Fig. 5.20(a). [O controlador é suposto um controlador proporcional do tipo indicado na Fig. 5.15(a).] Vamos supor que todas as variáveis r, q_i, h_1, e q_0 são medidas em relação a seus valores de regime estacionário $\bar{R}$, $\bar{Q}$, $\bar{H}$ e $\bar{Q}$. Suporemos também que as amplitudes das variáveis r, q_i, h_1 e q_0 sejam suficientemente pequenas de modo que o sistema possa ser aproximado por um modelo matemático linear, isto é, uma função de transferência.

Referindo-se à Seção 4.5, podemos obter a função de transferência do sistema de nível de líquido como

$$\frac{H_1(s)}{Q_i(s)} = \frac{R}{RCs + 1}$$

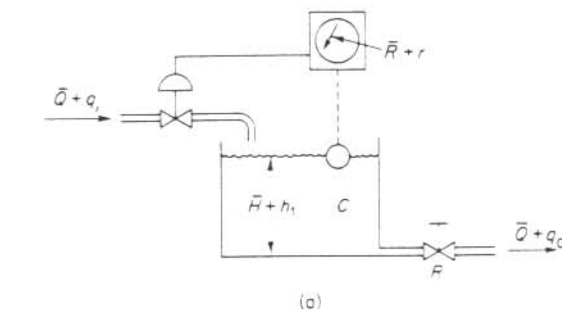
Desde que o controlador é um controlador proporcional, a variação no fluxo de entrada q_i é proporcional ao erro atuante e de modo que $q_i = K_p K_v e$, onde K_p é o ganho do controlador e K_v é o ganho da válvula de controle. Em termos das grandezas transformadas em Laplace,

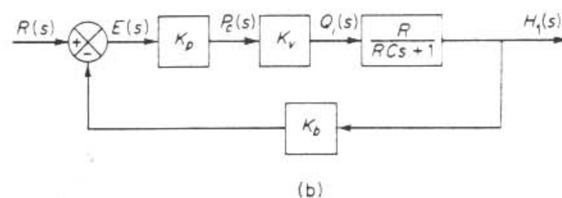
$$Q_i(s) = K_p K_v E(s)$$

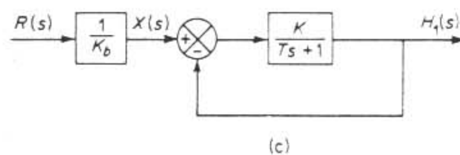
Um diagrama de blocos deste sistema é indicado na Fig. 5.20(b). Um diagrama de blocos simplificado é dado na Fig. 5.20(c), onde $X(s) = (1/K_b)R(s)$, $K = K_p K_v R K_b$ e $T = RC$.

A seguir investigaremos a resposta $h_1(t)$ a uma variação na entrada de referência. Suporemos uma variação em degrau unitário em $x(t)$, onde $x(t) = (1/K_b)r(t)$. A função de transferência de malha-fechada entre $H_1(s)$ e $X(s)$ é dada por

$$\frac{H_1(s)}{X(s)} = \frac{K}{Ts + 1 + K} \qquad (5\text{-}10)$$







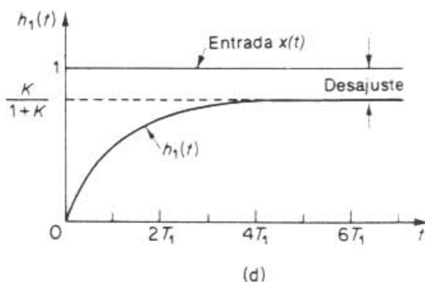


Fig. 5.20 (a) Sistema de controle de nível de líquido; (b) diagrama de blocos; (c) diagramas de blocos simplificado; (d) curva $h_1(t)$ *versus* t.

Desde que a transformada de Laplace da função degrau unitário é $1/s$, substituindo $X(s) = 1/s$ na Eq. (5.10), resulta

$$H_1(s) = \frac{K}{Ts + 1 + K}\frac{1}{s}$$

Expandindo $H_1(s)$ em frações parciais, obtemos

$$H_1(s) = \frac{K}{1+K}\frac{1}{s} - \frac{TK}{1+K}\frac{1}{Ts + 1 + K} \tag{5-11}$$

Tomando as transformadas de Laplace inversas dos dois membros da Eq. (5.11), obtemos a seguinte solução temporal $h_1(t)$:

$$h_1(t) = \frac{K}{1+K}(1 - e^{-t/T_1}) \qquad (t \geq 0) \tag{5-12}$$

onde

$$T_1 = \frac{T}{1+K}$$

A curva de resposta $h_1(t)$ é indicada na Fig. 5.20(d). A partir da Eq. (5.12), note que a constante de tempo T_1 do sistema de malha-fechada é diferente da constante de tempo T do bloco do ramo direto.

Da Eq. (5.12), verificamos que conforme t tende a infinito, o valor de $h_1(t)$ tende a $K/(1 + K)$, ou

$$h_1(\infty) = \frac{K}{1+K}$$

Desde que $x(\infty) = 1$, há um erro em regime estacionário de $1/(1 + K)$. Este erro é denominado desajuste (*offset*). O valor do desajuste torna-se menor conforme o ganho K se torna maior.

O desajuste é uma característica do controle proporcional de uma planta cuja função de transferência não possui um elemento integrante. (Obviamente, necessitamos de um erro não nulo de modo a possibilitar uma saída não nula.) Para eliminar este desajuste, devemos adicionar uma ação de controle integral. (Referir-se à Seção 5.3.)

Controladores proporcionais pneumáticos (tipo força-balanço). A Fig. 5.21 indica um diagrama esquemático de um controlador proporcional pneumático força-balanço. Controladores força-balanço são extensivamente usados na indústria. Estes controladores são muitas vezes denominados controladores de exaustão. O princípio básico de operação não difere daquele do controlador força-distância. A principal vantagem do controlador força-balanço é que ele elimina muitas perdas mecânicas e juntas por pivô, conseqüentemente reduzindo os efeitos de fricção.

A seguir, consideraremos o princípio do controlador força-balanço. No controlador indicado na Fig. 5.21, a pressão de entrada de referência P_r e a pressão de saída P_0 são injetadas em grandes câmaras por diafragma. Note que um controlador

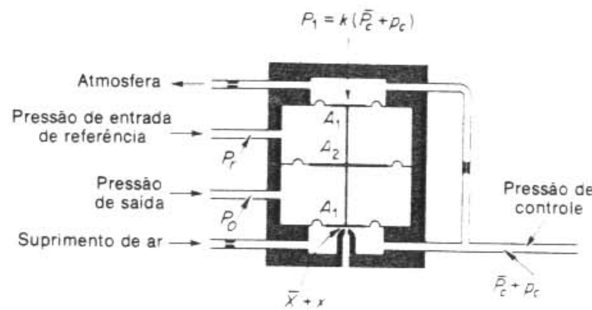


Fig. 5.21 Diagrama esquemático de um controlador pneumático proporcional com equilíbrio de força.

pneumático força-balanço opera apenas com sinais de pressão. Portanto, é necessário converter a entrada de referência e a saída do sistema em sinais de pressão correspondentes.

Como no caso do controlador força-distância, este controlador emprega uma haste, bocal e orifícios. Na Fig. 5.21, a abertura perfurada com broca, na câmara de baixo, é o bocal. O diafragma exatamente acima do bocal atua como uma "haste".

A operação do controlador força-balanço indicada na Fig. 5.21 pode ser sumariada como segue: 20 psig de ar de uma fonte de ar flui através de um orifício, causando uma pressão reduzida na câmara inferior. O ar nesta câmara escapa para a atmosfera através do bocal. O fluxo através do bocal depende tanto da abertura como da queda de pressão através dela. Um aumento na pressão de entrada de referência P_r, enquanto a pressão de saída P_0 se mantém a mesma, faz o pino da válvula mover-se para baixo, diminuindo a abertura entre o bocal e o diafragma "haste". Isto acarreta um acréscimo na pressão de controle P_c. Seja

$$p_e = P_r - P_0$$

Se $p_e = 0$, há um estado de equilíbrio com a distância bocal-haste igual a $\bar{X}$ e a pressão de controle igual a $\bar{P}_c$. Neste estado de equilíbrio, $P_1 = \bar{P}_c(k)$ (onde $k < 1$) e

$$\bar{X} = \alpha(\bar{P}_c A_1 - \bar{P}_c k A_1) \tag{5-13}$$

onde α é uma constante.

Vamos supor que $p_e \neq 0$ e definir pequenas variações na distância bocal-haste e pressão de controle como x e p_c, respectivamente. Obtemos então a seguinte equação:

$$\bar{X} + x = \alpha[(\bar{P}_c + p_c)A_1 - (\bar{P}_c + p_c)kA_1 - p_e(A_2 - A_1)] \tag{5-14}$$

Das Eqs. (5.13) e (5.14), obtemos

$$x = \alpha[p_c(1 - k)A_1 - p_e(A_2 - A_1)] \tag{5-15}$$

Neste ponto, devemos examinar a grandeza x. No projeto de controladores pneumáticos, a distância bocal-haste é muito pequena. Como x/α é um termo de ordem superior em relação a $p_c(1 - k)A_1$ ou $p_e(A_2 - A_1)$; isto é, para $p_e \neq 0$,

$$\frac{x}{\alpha} \ll p_c(1 - k)A_1$$

$$\frac{x}{\alpha} \ll p_e(A_2 - A_1)$$

devemos desprezar o termo x em nossa análise. A Eq. (5.15) pode então ser reescrita para refletir esta consideração, como segue:

$$p_c(1 - k)A_1 = p_e(A_2 - A_1)$$

e a função de transferência entre p_c e p_e resulta

$$\frac{P_c(s)}{P_e(s)} = \frac{A_2 - A_1}{A_1}\frac{1}{1 - k} = K_p$$

Conseqüentemente, o controlador indicado na Fig. 5.21 é um controlador proporcional. O valor do ganho K_p aumenta conforme k se aproxima da unidade. Note que o valor de k depende tanto dos orifícios da tubulação de entrada como da tubulação de saída da câmara de realimentação. (O valor de k tende à unidade conforme a resistência ao fluxo no orifício da tubulação de entrada é feito menor.)

Controladores proporcionais hidráulicos. Exceto para controladores pneumáticos de baixa pressão, ar comprimido raramente tem sido aplicado para controle contínuo do movimento de dispositivos possuindo massas significativas, sob ação de forças de carga externas. Para estes casos são normalmente preferidos os controladores hidráulicos. Os controladores hidráulicos também são extensivamente utilizados na indústria. Com sistemas hidráulicos de alta pressão podem ser obtidas forças muito grandes. É possível posicionamento preciso e rápido de grandes cargas com sistemas hidráulicos. Uma combinação de sistemas eletrônicos

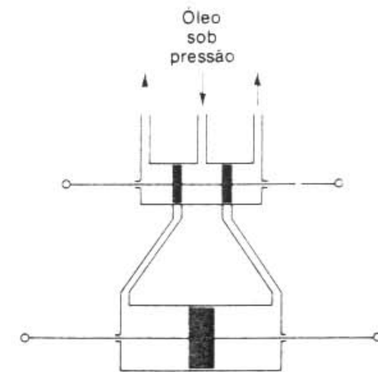


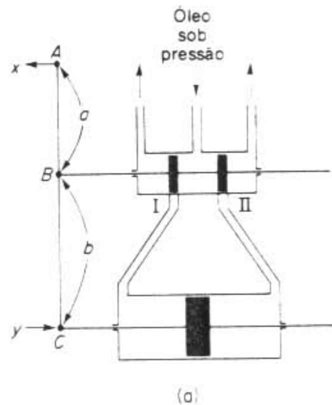
Fig. 5.22 Servomotor hidráulico.

e hidráulicos é muito utilizada porque combina as vantagens do controle eletrônico com a potência hidráulica.

Uma breve descrição da operação de um servomotor hidráulico foi fornecida na Seção 4.3. Foi mostrado que para massas de carga muito pequenas e desprezíveis, o servomotor indicado na Fig. 4.12 age como um integrador ou um controlador integral. [Referir-se à Eq. (4.25).] Este servomotor constitui a base do circuito de controle hidráulico.

O servomotor indicado na Fig. 5.22 age como um controlador integral. Podemos modificar este servomotor para um controlador proporcional incluindo um mecanismo de realimentação. A Fig. 5.23(a) é idêntica com a Fig. 5.22, exceto pelo elo atado ao lado esquerdo do elo de realimentação de junção do pistão de potência ABC em C. O elo AC é um elo flutuante ao inves de um elo móvel pelo pivô fixo. Veremos que o servomotor indicado na Fig. 5.23(a) age como um controlador proporcional.

Um sistema hidráulico indicado na Fig. 5.23(a) opera como segue: Se a entrada x move o pistão piloto para a esquerda, este abre o acesso I de modo que flui óleo a alta pressão através do acesso I do lado esquerdo do pistão de potência e força este pistão para a direita. O pistão de potência, movendo-se para a direita, leva o elo de realimentação AC com ele, movendo conseqüentemente o pistão piloto para a direita. Esta ação continua até que o pistão fecha novamente os acessos I e II. Um



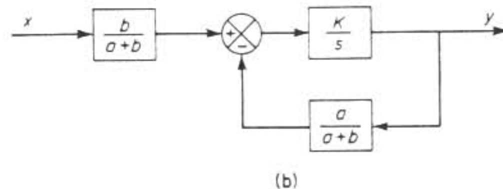


Fig. 5.23 (a) Servomotor que atua como um controlador proporcional; (b) diagrama de blocos do servomotor.

diagrama de blocos do sistema pode ser desenhado conforme a Fig. 5.23(b). A função de transferência entre y e x é dada por

$$\frac{Y(s)}{X(s)} = \frac{\dfrac{b}{a+b}\dfrac{K}{s}}{1 + \dfrac{K}{s}\dfrac{a}{a+b}} = \frac{bK}{s(a+b) + Ka}$$

Note que sob operação normal $Ka/s(a+b) \gg 1$, obtemos

$$\frac{Y(s)}{X(s)} = \frac{b}{a} = K_p$$

A função de transferência entre x e y resulta em uma constante, e portanto o sistema hidráulico mostrado na Fig. 5.23(a) age como um controlador proporcional, cujo ganho é K_p. O ganho K_p pode ser ajustado pelo ajuste efetivo da relação de alavanca b/a.

Verificamos que a adição de um braço de realimentação faz com que o servomotor hidráulico atue como um controlador proporcional.

Controladores proporcionais eletrônicos. Um controlador proporcional eletrônico é um amplificador que recebe um pequeno sinal de tensão e fornece uma saída

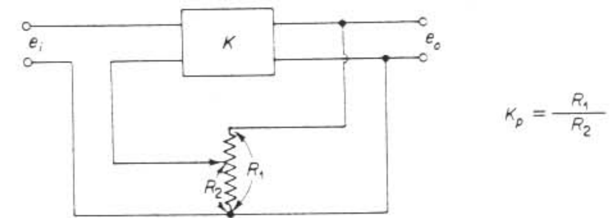


Fig. 5.24 Diagrama esquemático de um controlador proporcional eletrônico.

de tensão em um nível de potência maior. Um diagrama esquemático de um controlador deste tipo é mostrado na Fig. 5.24. Para este controlador,

$$e_0 = K\left(e_i - e_o\frac{R_2}{R_1}\right), \qquad K\frac{R_2}{R_1} \gg 1$$

Portanto, a função de transferência $G(s)$ deste controlador é

$$G(s) = \frac{E_0(s)}{E_i(s)} = \frac{R_1}{R_2} = K_p$$

K_p é o ganho do controlador proporcional. O ganho K_p pode ser ajustado variando-se a relação de resistências (R_1/R_2) no circuito de realimentação.

5.3 OBTENÇÃO DE UMA AÇÃO DE CONTROLE DERIVATIVA E INTEGRAL

Nesta seção apresentaremos métodos de obtenção de uma ação de controle derivativa e integral. Daremos novamente ênfase ao princípio e não a detalhes de mecanismos reais.

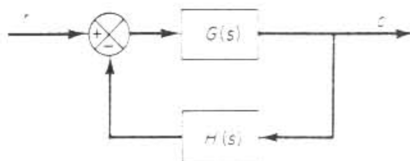


Fig. 5.25 Sistema de controle.

O princípio básico para gerar uma ação de controle desejada é inserir o inverso da função de transferência desejada no ramo de realimentação. Para o sistema indicado na Fig. 5.25, a função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)}$$

Se $|G(s)H(s)| \gg 1$, então $C(s)/R(s)$ pode ser modificado para

$$\frac{C(s)}{R(s)} = \frac{1}{H(s)}$$

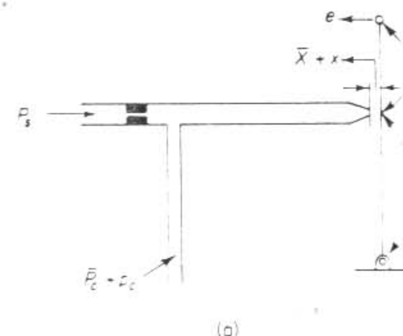
Portanto, se for desejada uma ação de controle proporcional-mais-derivativa, inserimos um elemento possuindo a função de transferência $1/(Ts + 1)$ no ramo de realimentação.

Obtenção de ação de controle proporcional-mais-derivativa pneumático. Considere o controlador pneumático indicado na Fig. 5.26(a). Este é um controlador proporcional de banda-estreita ou um controlador pneumático de duas posições. Suponha que o erro atuante nulo $e = 0$ corresponde à distância bocal-haste $\bar{X}$ e pressão de controle $\bar{P}_c$. Se uma pequena variação no erro atuante da posição zero produz uma pequena variação em x na distância bocal-haste, então uma pequena variação p_c é produzida na pressão de controle. Um diagrama de blocos do sistema (sob a hipótese de pequenas variações) pode ser determinado, conforme indicado na Fig. 5.26(b). A função de transferência entre p_c e e resulta

$$\frac{P_c(s)}{E(s)} = \frac{bK}{a + b} = K_p$$

Mostraremos a seguir que a adição de realimentação negativa atrasada ao controlador indicado na Fig. 5.26(a) modificará o controlador proporcional de banda-estreita para um controlador proporcional-mais-derivativo.

Considere o controlador pneumático indicado na Fig. 5.27(a). Supondo novamente pequenas variações no erro atuante, na distância bocal-haste, e na pressão de



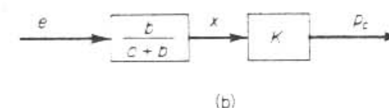


Fig. 5.26 (a) Um controlador proporcional pneumático; (b) diagrama de blocos do controlador.

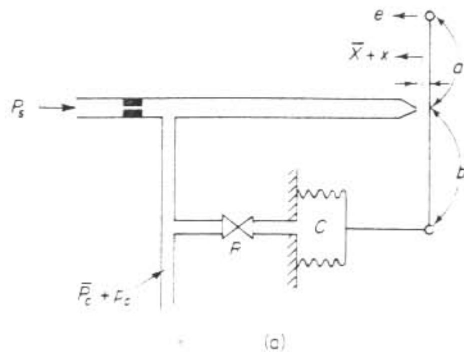
controle, podemos sumariar a operação deste controle como segue: Inicialmente suporemos uma pequena variação em degrau em e. Então, a variação na pressão de controle p_c será instantânea. A restrição R evitará que o fole de realimentação sinta a variação de pressão p_c. Portanto o fole de realimentação não responderá momentaneamente, e a válvula atuante pneumática receberá o efeito total do movimento da haste. Com o passar do tempo o fole de realimentação se expande ou se contrai. A variação na distância bocal-haste x e a variação na pressão de controle p_c podem ser desenhadas em um gráfico em função do tempo t, como indicado na Fig. 5.27(b). Em regime estacionário, o fole de realimentação atua como um mecanismo de realimentação comum. A curva p_c *versus* t mostra claramene que este controlador é o tipo proporcional-mais-derivativo.

Um diagrama de blocos correspondente a este controlador pneumático é indicado na Fig. 5.27(c). No diagrama de blocos, K é uma constante, A é a área de fole e k_s é a constante de mola equivalente do fole. A função de transferência entre p_c e e pode ser obtida do diagrama de blocos, resultando:

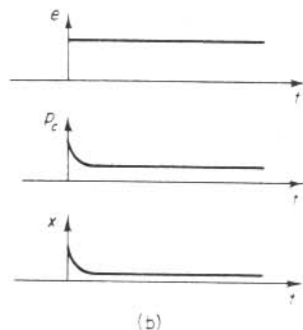
$$\frac{P_c(s)}{E(s)} = \frac{\dfrac{b}{a + b}K}{1 + \dfrac{Ka}{a + b}\dfrac{A}{k_s}\dfrac{1}{RCs + 1}}$$

Neste tipo de controlador o ganho de laço $|KaA/[(a + b)k_s(RCs + 1)]|$ normalmente é muito maior do que 1. Conseqüentemente a função de transferência $P_c(s)/E(s)$ pode ser simplificada para

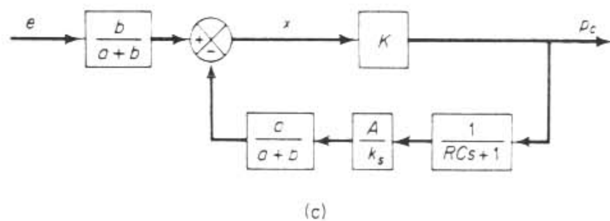
$$\frac{P_c(s)}{E(s)} = K_p(1 + T_d s)$$



(a)



(b)



(c)

Fig. 5.27 (a) Um controlador pneumático proporcional-mais-derivativo; (b) variação em degrau em e e as variações correspondentes em p_c e x em um gráfico *versus t;* (c) diagrama de blocos do controlador.

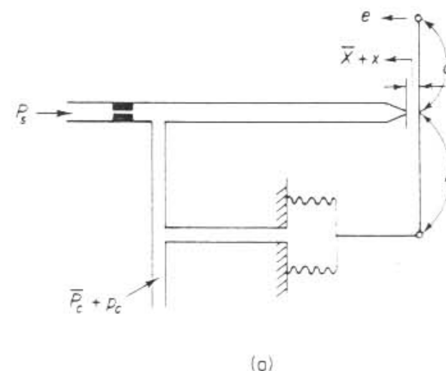
onde

$$K_p = \frac{bk_s}{aA}, \qquad T_d = RC$$

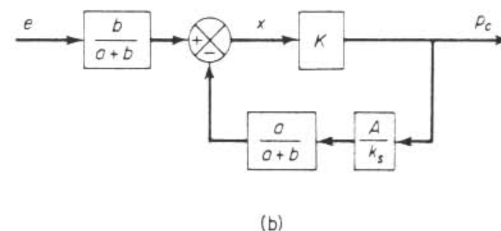
Portanto, a realimentação negativa atrasada ou a função de transferência $1/(RCs + 1)$ no ramo de realimentação modifica o controlador proporcional para um controlador proporcional-mais-derivativo.

Note que se a válvula de realimentação estiver completamente aberta, resulta uma ação de controle proporcional. Se a válvula de realimentação estiver completamente fechada, resulta uma ação de controle proporcional com banda-estreita (liga-desliga).

Obtenção da ação de controle proporcional-mais-integral pneumática. Considere o controlador proporcional indicado na Fig. 5.28(a). Considerando pequenas variações nas variáveis, podemos desenhar um diagrama de blocos deste controlador conforme a Fig. 5.28(b). Mostraremos que a adição de realimentação positiva atrasada modificará este controlador proporcional para um controlador proporcional-mais-integral.



(a)



(b)

Fig. 5.28 Um controlador proporcional pneumático; (b) diagrama de blocos do controlador.

Considere o controlador pneumático indicado na Fig. 5.29(a). A operação deste controlador é a seguinte: O fole indicado por I é ligado à fonte de pressão de controle sem qualquer restrição. O fole indicado por II é ligado à fonte de pressão de controle através de uma restrição. Vamos supor uma pequena variação no erro atuante. Isto acarretará uma variação instantânea na pressão do bocal. Portanto, uma variação também ocorrerá instantaneamente na pressão de controle p_c. Devido à restrição da válvula no caminho para o fole II, haverá uma queda de pressão na válvula. Com o transcorrer do tempo, fluirá ar através da válvula de tal modo que a variação na pressão do fole II atingirá o valor p_c. Conseqüentemente, o fole II se expandirá ou contrairá conforme transcorre o tempo, de tal modo a mover a haste de uma quantidade adicional no sentido do deslocamento original e. Isto acarretará uma variação contínua na pressão p_c do bocal, conforme indicado na Fig. 5.29(b).

Note que a ação de controle integral no controlador toma a forma de um lento cancelamento na realimentação fornecida originalmente pelo controle proporcional.

Um diagrama de blocos deste controlador sob a hipótese de pequenas variações nas variáveis é indicado na Fig. 5.29(c). Uma simplificação deste diagrama de blocos resulta na Fig. 5.29(d). A função de transferência deste controlador é

$$\frac{P_c(s)}{E(s)} = \frac{\frac{b}{a+b}K}{1 + \frac{Ka}{a+b}\frac{A}{k_s}\left(1 - \frac{1}{RCs+1}\right)}$$

onde K é uma constante, A a área dos foles e k_s a constante de mola equivalente dos foles combinados. Se $KaARCs/[(a+b)k_s(RCs+1)] \gg 1$, que é normalmente o caso, a função de transferência pode ser simplificada para

$$\frac{P_c(s)}{E(s)} = K_p\left(1 + \frac{1}{T_i s}\right)$$

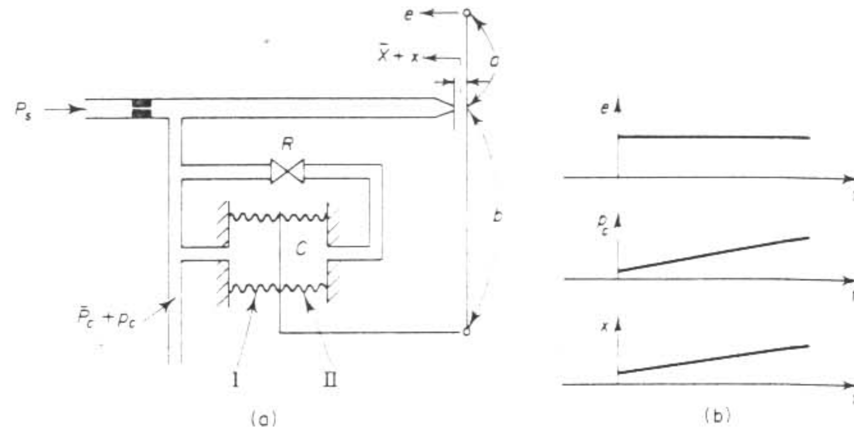
onde

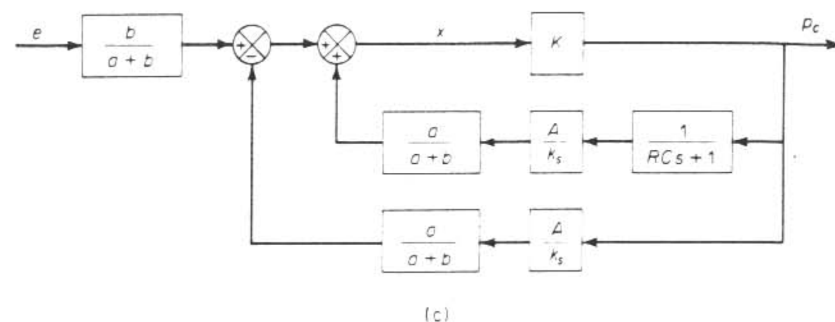
$$K_p = \frac{bk_s}{aA}, \qquad T_i = RC$$

Obtenção de ação de controle proporcional-mais-derivativa-mais-integral pneumática. Uma combinação das ações de controle pneumáticas indicadas nos sistemas das Figs. 5.27(a) e 5.29(a) fornece uma ação de controle proporcional-mais-derivativa-mais-integral. A Fig. 5.30(a) mostra um diagrama esquemático deste tipo de controlador. A Fig. 5.30(b) mostra o diagrama de blocos deste controlador sob a hipótese de pequenas variações nas variáveis.

A função de transferência deste controlador é

$$\frac{P_c(s)}{E(s)} = \frac{\frac{bK}{a+b}}{1 + \frac{Ka}{a+b}\frac{A}{k_s}\frac{(R_iC - R_dC)s}{(R_dCs+1)(R_iCs+1)}}$$





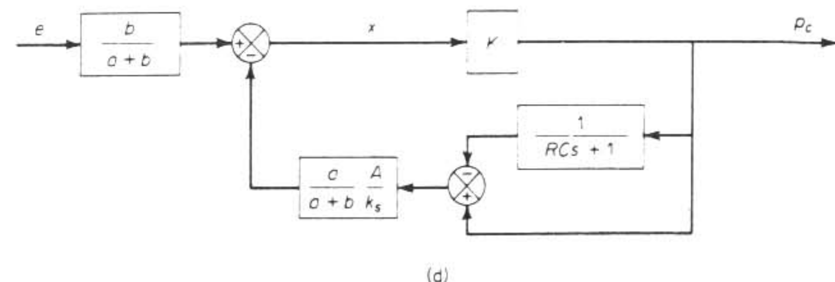


Fig. 5.29 (a) Um controlador pneumático proporcional-mais-integral; (b) variação em degrau em e e variações correspondentes em p_c e x em função de t; (c) diagramas de blocos do controlador; (d) diagrama de bloco simplificado.

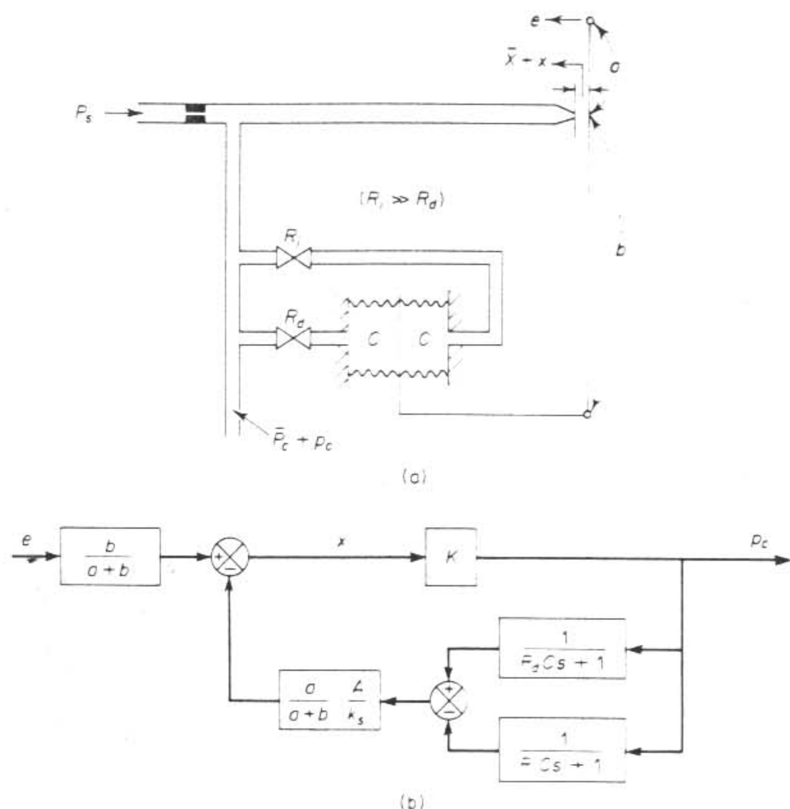


Fig. 5.30 (a) Um controlador pneumático proporcional-mais-derivativo-mais-integral; (b) diagramas de blocos do controlador.

Definindo,

$$T_i = R_i C, \qquad T_d = R_d C$$

e considerando que sob operação normal $|KaA(T_i - T_d)s/[(a+b)k_s(T_d s + 1)(T_i s + 1)]| >> 1$ e $T_i >> T_d$, obtemos

$$\frac{P_c(s)}{E(s)} \doteq \frac{bk_s}{aA} \frac{(T_d s + 1)(T_i s + 1)}{(T_i - T_d)s}$$

$$\doteq \frac{bk_s}{aA} \frac{T_d T_i s^2 + T_i s + 1}{T_i s}$$

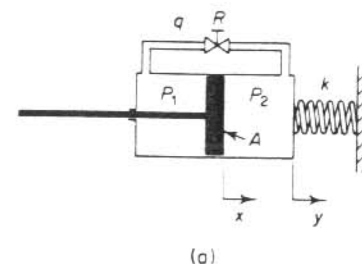
$$= K_p\left(1 + T_d s + \frac{1}{T_i s}\right) \qquad (5\text{-}16)$$

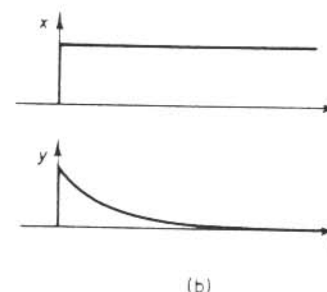
onde

$$K_p = \frac{bk_s}{aA}$$

A Eq. (5.16) indica que o controlador mostrado na Fig. 5.30(a) é um controlador proporcional-mais-derivativo-mais-integral.

Amortecedores viscosos. O amortecedor viscoso indicado na Fig. 5.31(a) atua como um elemento diferenciador. Suponha que introduzimos um deslocamento integral na posição x do pistão. Então, o deslocamento y, momentaneamente, torna-se igual a x. Devido à força da mola, entretanto, o óleo fluirá através da





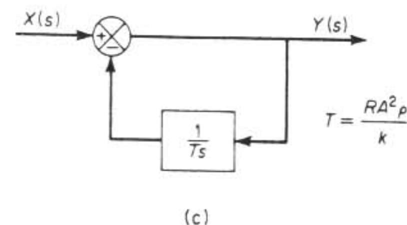


Fig. 5.31 (a) Amortecedor viscoso; (b) variação em degrau em x e variação correspondente em y, em função de t; (c) diagrama de blocos do amortecedor viscoso.

resistência R e o cilindro retornará à posição original. As curvas x *versus* t e y *versus* t são indicadas na Fig. 5.31(b).

Para definir a função de transferência entre o deslocamento y e o deslocamento x, vamos definir as pressões existentes nos lados direito e esquerdo do pistão como P_1 (N/m²) e P_2 (N/m²), respectivamente. Suponha que a força de inércia envolvida seja desprezível. Então, a força atuante no pistão deve equilibrar a força da mola. Portanto,

$$A(P_2 - P_1) = ky$$

onde

A = área do pistão, m²
k = constante da mola, N/m.

A taxa de fluxo q é dada por

$$q = \frac{P_2 - P_1}{R}$$

onde

q = taxa de fluxo através da restrição, kg/s
R = resistência ao fluxo na restrição, N-s/m²-kg

Desde que o fluxo através da restrição durante dt segundos deve ser igual à variação da massa de óleo para a esquerda do pistão durante os mesmos dt segundos, obtemos

$$q\,dt = A\rho(dx - dy)$$

onde

ρ = densidade, kg/m³

(Supusemos que o fluido é incompreensível ou ρ = constante.) Esta última equação pode ser reescrita como

$$\frac{dx}{dt} - \frac{dy}{dt} = \frac{q}{A\rho} = \frac{P_2 - P_1}{RA\rho} = \frac{ky}{RA^2\rho}$$

ou

$$\frac{dx}{dt} = \frac{dy}{dt} + \frac{ky}{RA^2\rho}$$

Considerando as transformadas de Laplace de ambos os membros desta última equação, supondo condições iniciais nulas, obtemos

$$sX(s) = sY(s) + \frac{k}{RA^2\rho}Y(s)$$

A função de transferência deste sistema resulta

$$\frac{Y(s)}{X(s)} = \frac{s}{s + \dfrac{k}{RA^2\rho}}$$

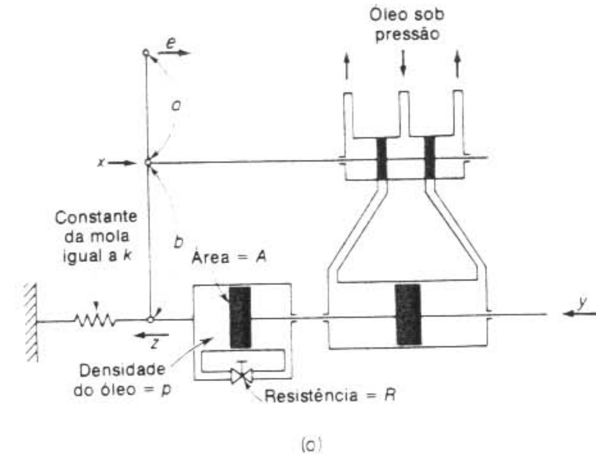
Vamos definir $RA^2\rho/k = T$. Portanto,

$$\frac{Y(s)}{X(s)} = \frac{Ts}{Ts + 1} = \frac{1}{1 + \dfrac{1}{Ts}}$$

A Fig. 5.31(c) mostra uma representação em diagrama de blocos para este sistema.

Obtenção da ação de controle proporcional-mais-integral hidráulica. A Fig. 5.32(a) mostra um diagrama esquemático de um controlador proporcional-mais-integral hidráulico. Um diagrama de blocos deste controlador é indicado na Fig. 5.32(b). A função de transferência $Y(s)/E(s)$ é dada por

$$\frac{Y(s)}{E(s)} = \frac{\dfrac{b}{a+b}\dfrac{K}{s}}{1 + \dfrac{Ka}{a+b}\dfrac{T}{Ts+1}}$$



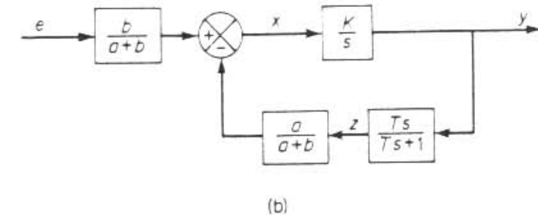


Fig. 5.32 (a) Diagrama esquemático de um controlador proporcional-mais-integral; (b) diagrama de blocos do controlador.

Neste controlador, sob operação normal $|KaT/[(a+b)(Ts+1)]| \gg 1$, resultando

$$\frac{Y(s)}{E(s)} = K_p\left(1 + \frac{1}{T_i s}\right)$$

onde

$$K_p = \frac{b}{a}, \qquad T_i = T = \frac{RA^2\rho}{k}$$

Portanto o controlador indicado na Fig. 5.32(a) é um controlador proporcional-mais-integral.

Obtenção de ação de controle derivativa e integral em controladores eletrônicos. A Fig. 5.33 indica o princípio de obtenção da ação de controle derivativa e integral em controladores eletrônicos. Essencialmente, introduzimos um circuito apropriado no ramo de realimentação a fim de gerar a ação de controle desejada. As funções de transferência dos controladores podem ser obtidas como segue: Para o controlador indicado na Fig. 5.33(a),

$$\frac{E_f(s)}{E_0(s)} = \frac{1}{R_d C_d s + 1}$$

$$[E_i(s) - E_f(s)]K = E_0(s)$$

Portanto, para $|K/(R_d C_d s + 1)| \gg 1$,

$$\frac{E_0(s)}{E_i(s)} = \frac{K(R_d C_d s + 1)}{R_d C_d s + 1 + K} \doteq R_d C_d s + 1 = T_d s + 1$$

onde $T_d = R_d C_d$.

Analogamente, para o controlador indicado na Fig. 5.33(b),

$$\frac{E_f(s)}{E_0(s)} = \frac{R_i C_i s}{R_i C_i s + 1}$$

$$[E_i(s) - E_f(s)]K = E_0(s)$$

Portanto, para $|KR_i C_i s/(R_i C_i s + 1)| \gg 1$,

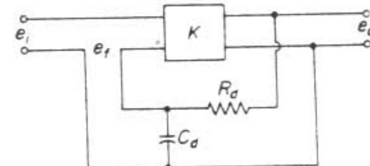
$$\frac{E_0(s)}{E_i(s)} = \frac{K(R_i C_i s + 1)}{KR_i C_i s + R_i C_i s + 1} \doteq \frac{R_i C_i s + 1}{R_i C_i s} = 1 + \frac{1}{T_i s}$$

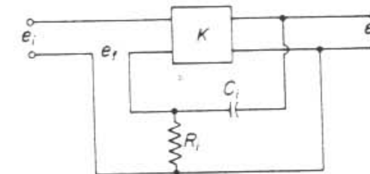
onde

$$T_i = R_i C_i.$$

Para o controlador indicado na Fig. 5.33(c), se o ganho de laço é muito maior do que a unidade, a função de transferência pode ser obtida como

$$\frac{E_0(s)}{E_i(s)} = K_p\alpha\left(1 + \frac{T_d}{\alpha}s + \frac{1}{\alpha T_i s}\right)$$





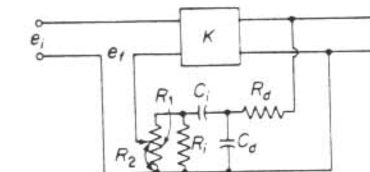


Fig. 5.33 Controladores eletrônicos. (a) controlador proporcional-mais-derivativo; (b) controlador proporcional-mais-integral; (c) controlador proporcional-mais-derivativo-mais-integral.

onde

$$\alpha = 1 + \frac{R_d}{R_i} + \frac{T_d}{T_i}$$

A dedução desta função de transferência é fornecida no Problema A.5.5.

5.4 EFEITOS DA AÇÃO DE CONTROLE INTEGRAL E DERIVATIVA NO DESEMPENHO DO SISTEMA

Nesta seção investigaremos os efeitos da ação de controle integral e derivativa no desempenho do sistema, porém consideraremos apenas sistemas simples. (Nos últimos capítulos faremos um estudo adicional sobre controle integral e derivativo.)

Ação de controle integral. No controle proporcional de um processo cuja função de transferência não possui um integrador $1/s$, há um erro em regime estacionário, ou desajuste, na resposta à entrada ao degrau. Este desajuste pode ser eliminado se for incluída no controlador uma ação de controle integral.

No controle integral de um processo, o sinal de controle, o sinal de saída do controlador, em qualquer instante, é igual à área sob a curva do sinal erro atuante até aquele instante. O sinal de controle $m(t)$ pode possuir um valor não nulo quando o sinal de erro atuante $e(t)$ é nulo, conforme indicado na Fig. 5.34(a). Isto é impossível no caso do controlador proporcional desde que um sinal de controle não nulo necessita de um sinal de erro atuante não nulo. (Um sinal de erro atuante não nulo em regime estacionário significa que há um desajuste.) A Fig. 5.34(b) mostra a curva $e(t)$ *versus* t e a curva correspondente $m(t)$ *versus* t quando o controlador é do tipo proporcional.

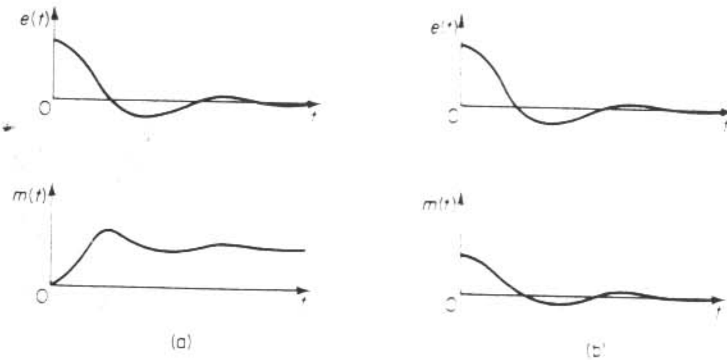


Fig. 5.34 (a) Curvas de $e(t)$ e $m(t)$ indicando sinal de controle não nulo quando o sinal erro atuante é nulo; (b) curvas de $e(t)$ e $m(t)$ mostrando sinal de controle nulo quando o sinal de erro atuante é nulo.

Note que a ação de controle integral, embora remova o desajuste ou erro em regime estacionário, pode resultar em uma resposta oscilatória com amplitude ligeiramente decrescente ou mesmo com amplitude crescente, ambas usualmente indesejáveis. (Para detalhes, vide Cap. 6.)

Controle integral de sistemas de controle de nível de líquido. Na Seção 5.2, verificamos que o controle proporcional de um sistema de nível de líquido resulta em um erro em regime estacionário com uma entrada em degrau. Mostraremos agora que este erro pode ser eliminado se for incluída no controlador uma ação de controle integral.

A Fig. 5.35(a) mostra um sistema de controle de nível de líquido. Suporemos que o controlador é um controlador integral. Consideraremos ainda que as variáveis x, q_i, h e q_0, medidas em relação aos seus valores respectivos de regime estacionário $\bar{X}$, $\bar{Q}$, $\bar{H}$ e $\bar{Q}$, são pequenas quantidades de modo que o sistema pode ser considerado linear. Sob estas hipóteses, o diagrama de blocos do sistema pode ser obtido como indicado na Fig. 5.35(b). Da Fig. 5.35(b), a função de transferência de malha-fechada entre $H(s)$ e $X(s)$ é

$$\frac{H(s)}{X(s)} = \frac{KR}{RCs^2 + s + KR}$$

Portanto,

$$\frac{E(s)}{X(s)} = \frac{X(s) - H(s)}{X(s)} = \frac{RCs^2 + s}{RCs^2 + s + KR}$$

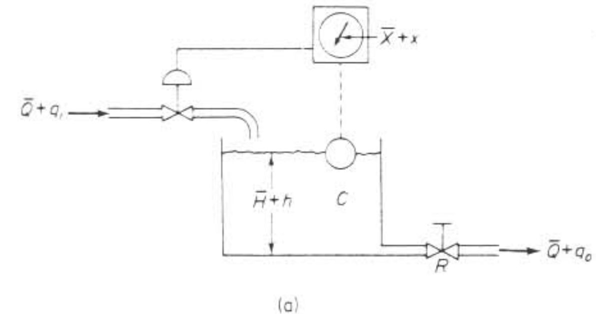


Fig. 5.35 (a) Sistema de controle de nível de líquido; (b) diagrama de blocos do sistema.

Desde que o sistema é estável, o erro em regime estacionário para a resposta em degrau unitário é obtido pela aplicação do teorema do valor final como segue:

$$e_{ss} = \lim_{s\to 0} sE(s) = \lim_{s\to 0} \frac{s(RCs^2 + s)}{RCs^2 + s + KR}\frac{1}{s} = 0$$

O controle integral do sistema de nível de líquido, portanto, elimina o erro em regime estacionário na resposta à entrada em degrau. Esta é uma vantagem importante em relação ao controle proporcional que resulta em desajuste.

Resposta para distúrbios de torque (controle proporcional). Vamos investigar o efeito de um distúrbio de torque que ocorre em um elemento de carga. Considere o sistema indicado na Fig. 5.36. O controlador proporcional fornece um torque T para posicionar o elemento de carga, que consiste em momento de inércia e atrito viscoso. O distúrbio de torque é denotado por N.

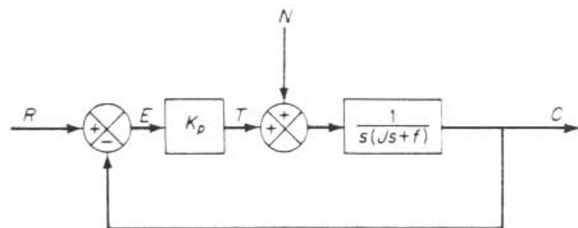


Fig. 5.36 Sistema de controle com um distúrbio de torque.

Supondo nula a entrada de referência ou $R(s) = 0$, a função de transferência entre $C(s)$ e $N(s)$ é dada por

$$\frac{C(s)}{N(s)} = \frac{1}{Js^2 + fs + K_p}$$

Portanto,

$$\frac{E(s)}{N(s)} = -\frac{C(s)}{N(s)} = -\frac{1}{Js^2 + fs + K_p}$$

O erro em regime estacionário devido a um distúrbio de torque em degrau de valor T_n é dado por

$$\begin{aligned} e_{ss} &= \lim_{s\to 0} sE(s) \\ &= \lim_{s\to 0} \frac{-s}{Js^2 + fs + K_p}\frac{T_n}{s} \\ &= -\frac{T_n}{K_p} \end{aligned}$$

Em regime permanente, o controlador proporcional fornece o torque $-T_n$, que é igual em nódulo, mas oposto em sinal, ao torque de distúrbio T_n. A saída em regime permanente, devida ao torque de distúrbio em degrau, é

$$c_{ss} = -e_{ss} = \frac{T_n}{K_p}$$

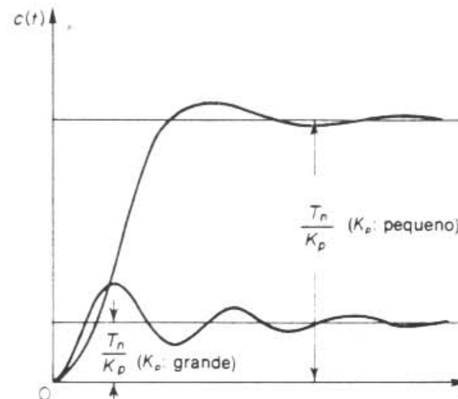


Fig. 5.37 Curvas de resposta típicas para o distúrbio de torque em degrau.

O erro em regime estacionário pode ser reduzido pelo aumento do valor do ganho K_p. O aumento deste valor, entretanto, resultará em uma resposta mais oscilatória do sistema. Curvas de resposta típicas para um pequeno valor de K_p e um grande valor de K_p são indicadas na Fig. 5.37.

Desde que o valor do ganho K_p não pode ser aumentado demasiadamente, é desejável modificar o controlador proporcional para um controlador proporcional-mais-integral.

Resposta a distúrbios de torque (controle proporcional-mais-integral). Para eliminar o desajuste devido ao distúrbio de torque, o controlador proporcional deve ser substituído por um controlador proporcional-mais-integral.

Se for adicionada ao controlador uma ação de controle integral, enquanto houver um sinal de erro, haverá um torque desenvolvido pelo controlador para reduzir este erro, desde que o sistema de controle é estável.

A Fig. 5.38 mostra o controle proporcional-mais-integral do elemento de carga, consistindo em momento de inércia e atrito viscoso.

A função de transferência de malha-fechada entre $C(s)$ e $N(s)$ é

$$\frac{C(s)}{N(s)} = \frac{s}{Js^3 + fs^2 + K_p s + \dfrac{K_p}{T_i}}$$

Na ausência da entrada de referência, ou $r(t) = 0$, o sinal erro é obtido de

$$E(s) = -\frac{s}{Js^3 + fs^2 + K_p s + \dfrac{K_p}{T_i}}N(s)$$

Se este sistema de controle for estável, isto é, se as raízes da equação característica

$$Js^3 + fs^2 + K_p s + \frac{K_p}{T_i} = 0$$

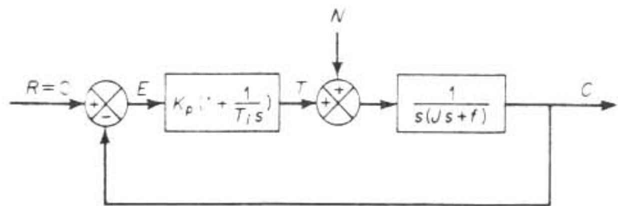


Fig. 5.38 Controle proporcional-mais-integral de um elemento de carga consistindo em momento de inércia e fricção viscosa.

possuem partes reais negativas, então o erro em regime estacionário na resposta a um torque de distúrbio em degrau de valor T_n é obtido aplicando-se o teorema de valor final como segue:

$$e_{ss} = \lim_{s\to 0} sE(s)$$

$$= \lim_{s\to 0} \frac{-s^2}{Js^3 + fs^2 + K_p s + \dfrac{K_p}{T_i}} \frac{T_n}{s}$$

$$= 0$$

Portanto o erro em regime estacionário em relação ao distúrbio por torque pode ser eliminado se o controlador for do tipo proporcional-mais-integral.

Note que a ação de controle integral adicionada ao controlador proporcional converteu o sistema originalmente de segunda-ordem em um sistema de terceira-ordem. Conseqüentemente, o sistema de controle pode tornar-se instável para um valor grande de K_p desde que as raízes da equação característica podem possuir partes reais positivas. (O sistema de segunda-ordem é sempre estável se os coeficientes na equação diferencial do sistema são todos positivos.)

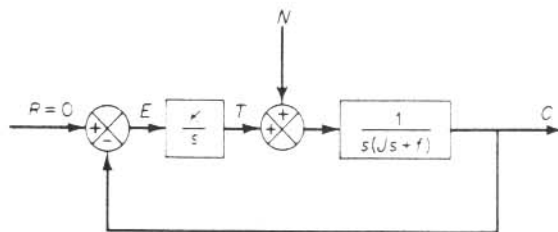


Fig. 5.39 Controle integral de um elemento de carga consistindo em momento de inércia e fricção viscosa.

É importante observar que, se o controlador fosse um controlador integral, como na Fig. 5.39, então o sistema sempre resultaria instável porque a equação característica

$$Js^3 + fs^2 + K = 0$$

possuiria raízes com partes reais positivas. Um sistema instável deste tipo não poderia ser usado na prática.

Note que, no sistema da Fig. 5.38, a ação de controle proporcional tende a estabilizar o sistema, enquanto a ação de controle integral tende a eliminar ou reduzir o erro em regime estacionário em resposta a várias entradas.

Ação de controle derivativa. A ação de controle derivativa, quando adicionada a um controlador proporcional, possibilita um meio de obter um controlador com alta sensibilidade. Uma vantagem em usar ação de controle derivativa é que ela responde à taxa de variação do erro atuante e pode produzir uma correção significativa antes de o valor do erro atuante tornar-se demasiadamente grande. O controle derivativo, portanto, antecipa o erro atuante e inicia uma ação corretiva mais cedo, tendendo a aumentar a estabilidade do sistema.

Embora o controle derivativo não afete diretamente o erro em regime estacionário, ele introduz amortecimento no sistema e portanto permite o uso de um valor maior do ganho K, o que resulta em uma melhora na precisão em regime estacionário.

Devido ao fato de o controle derivativo operar sobre a taxa de variação do erro atuante e não sobre o próprio erro atuante, este modo nunca é usado sozinho. É sempre utilizado em combinação com ação proporcional ou ação proporcional-mais-integral.

Sistemas de controle proporcional com carga de inércia. Antes de discutirmos o efeito da ação derivativa no desempenho do sistema, consideraremos o controle proporcional de uma carga de inércia.

Considere o sistema indicado na Fig. 5.40(a). A função de transferência de malha-fechada é obtida como

$$\frac{C(s)}{R(s)} = \frac{K_p}{Js^2 + K_p}$$

Desde que as raízes da equação característica

$$Js^2 + K_p = 0$$

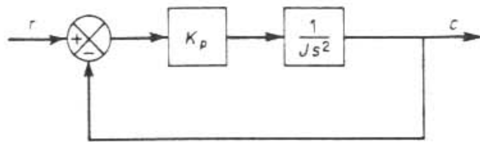
são imaginárias, a resposta a uma entrada em degrau unitário continua a oscilar indefinidamente, como indicado na Fig. 5.40(b).

Sistemas de controle exibindo estas características de resposta não são desejáveis. Veremos que a adição de um controle derivativo estabilizará o sistema.

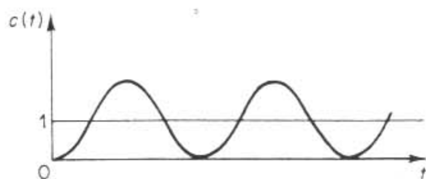
Controle proporcional-mais-derivativo de um sistema com carga de inércia. Modifiquemos o controlador proporcional para um controlador proporcional-mais-derivativo cuja função de transferência é $K_p(1 + T_d s)$. O torque desenvolvido pelo controlador é proporcional a $K_p(e + T_d\dot{e})$. O controle derivativo é essencialmente antecipatório, mede a velocidade de erro intantânea, prediz a grande sobrelevação antecipadamente no tempo e produz uma ação contrária apropriada antes de ocorrer uma sobrelevação demasiadamente grande.

Considere o sistema indicado na Fig. 5.41(a). A função de transferência de malha-fechada é dada por

$$\frac{C(s)}{R(s)} = \frac{K_p(1 + T_d s)}{Js^2 + K_p T_d s + K_p}$$

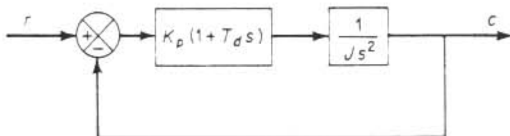


(a)

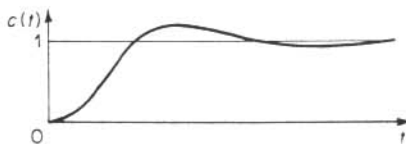


(b)

Fig. 5.40 (a) Controle proporcional de um sistema com carga de inércia; (b) resposta a uma entrada em degrau unitário.



(a)



(b)

Fig. 5.41 (a) Controle proporcional-mais-derivativo com um sistema com carga de inércia; (b) resposta a uma entrada em degrau unitário.

A equação característica

$$Js^2 + K_pT_ds + K_p = 0$$

agora possui duas raízes com partes reais negativas para valores positivos de J, K_p e T_d. Portanto, o controle derivativo introduz um efeito de amortecimento. Uma curva de resposta típica $c(t)$ para uma entrada com degrau unitário é fornecida na Fig. 5.41(b). Obviamente, a curva de resposta mostra uma melhoria significativa em relação à curva de resposta original indicada na Fig. 5.40(b).

5.5 REDUÇÃO DE VARIAÇÕES DE PARÂMETROS PELO USO DE REALIMENTAÇÃO

A principal finalidade de usar realimentação em sistemas de controle é reduzir a sensibilidade do sistema em relação a variações de parâmetros e distúrbios indesejáveis.

Se necessitarmos construir um sistema de controle de malha-aberta conveniente, devemos selecionar todos os componentes da função de transferência de malha-aberta $G(s)$ muito cuidadosamente de modo que respondam precisamente. No caso de construir um sistema de controle de malha-fechada, entretanto, os componentes podem ser menos precisos desde que a sensibilidade em relação a variações de parâmetros em $G(s)$ é reduzida por um fator de $1 + G(s)$.

Para ilustrar esta afirmação, considere os sistemas de malha-aberta e de malha-fechada indicados nas Figs. 5.42(a) e (b), respectivamente. Suponha que devido a variações de parâmetros, $G(s)$ é variada para $G(s) + \Delta G(s)$, onde $|G(s)| >> |\Delta G(s)|$. Então, no sistema de malha-aberta indicado na Fig. 5.42(a), a saída é dada por

$$C(s) + \Delta C(s) = [G(s) + \Delta G(s)]R(s)$$

Portanto, a variação na saída é dada por

$$\Delta C(s) = \Delta G(s)R(s)$$

No sistema de malha-fechada indicado na Fig. 5.42(b),

$$C(s) + \Delta C(s) = \frac{G(s) + \Delta G(s)}{1 + G(s) + \Delta G(s)}R(s)$$

ou

$$\Delta C(s) \doteq \frac{\Delta G(s)}{1 + G(s)}R(s)$$

Portanto, a variação na saída do sistema de malha-fechada, devida a variações de parâmetros em $G(s)$, é reduzida por um fator $1 + G(s)$. Em muitos casos práticos, o módulo de $1 + G(s)$ é normalmente muito maior do que um.

Note que, reduzindo os efeitos das variações dos parâmetros dos componentes, estamos também contornando o componente que causa a variação, com um laço de realimentação.

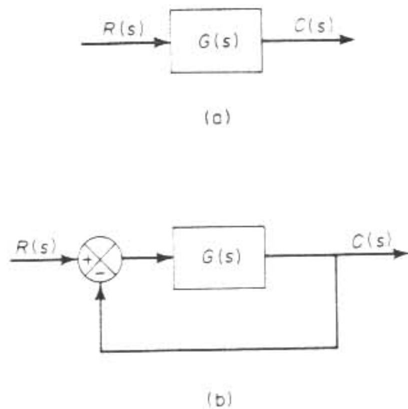


Fig. 5.42 (a) Sistema de malha-aberta; (b) sistema de malha-fechada.

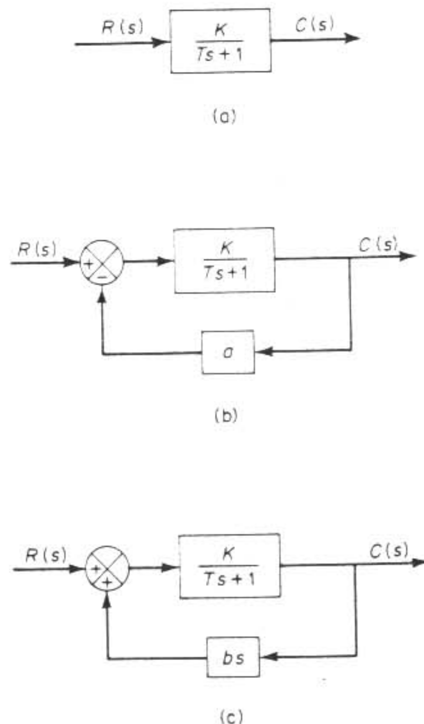


Fig. 5.43 (a) Sistema de malha-aberta; (b) sistema de malha-fechada com constante de tempo $T/(1 + Ka)$; (c) sistema de malha-fechada com constante de tempo $T - bK$.

Modificações de constante de tempo pelo uso de realimentação. Considere o sistema mostrado na Fig. 5.43(a). A constante de tempo do sistema é T. A adição de uma malha de realimentação negativa em volta deste elemento reduz a constante de tempo. A Fig. 5.43(b) mostra o sistema com a mesma função de transferência do ramo direto daquela indicada na Fig. 5.43(a), com a exceção da malha de realimentação negativa que foi adicionada. A constante de tempo deste sistema foi reduzida para $T/(1 + Ka)$. Note também que a constante de ganho para este sistema foi reduzida de K para $K/(1 + Ka)$.

Se, ao invés de uma malha de realimentação negativa, fosse adicionada uma malha de realimentação positiva em volta da função de transferência $K/(Ts + 1)$ e se a função de transferência da realimentação for escolhida apropriadamente, então a constante de tempo seria nula ou com um valor muito pequeno. Considere o sistema indicado na Fig. 5.43(c). Desde que a função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{K}{(T - bK)s + 1}$$

a constante de tempo pode ser reduzida através de uma escolha apropriada de b. Se b é escolhida igual a T/K, então resulta uma constante de tempo nula. Note, entretanto, que se os distúrbios acarretam um valor de $T - bK$ negativo ao invés de nulo, resulta um sistema instável. Então, se for empregada realimentação positiva para reduzir a constante de tempo para um valor pequeno, devemos tomar cuidado para que $T - bK$ nunca se torne negativo.

Aumento de ganhos de malha pelo uso de realimentação positiva. O sistema indicado na Fig. 5.44(a) possui a função de transferência $C(s)/R(s) = G(s)$. Consi-

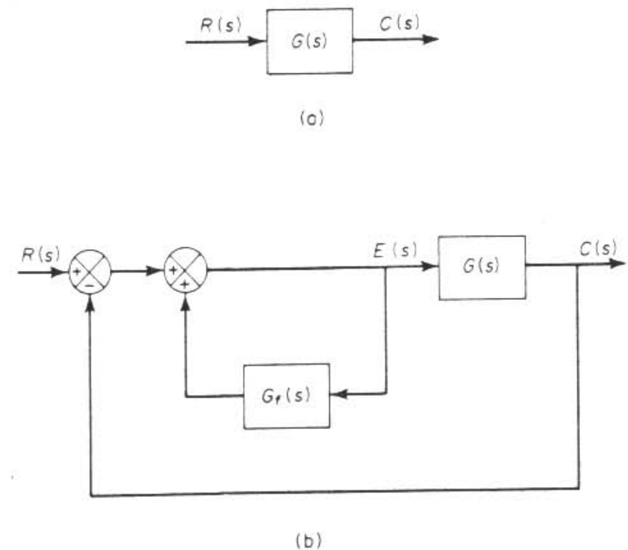


Fig. 5.44 (a) Sistema de malha-aberta; (b) sistema de malha-fechada cuja função de transferência é aproximadamente unitária.

dere agora o sistema indicado na Fig. 5.44(b). A função de transferência de malha-fechada para este sistema é

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 - G_f(s) + G(s)}$$

Se $G_f(s)$ é escolhida próxima à unidade, ou $G_f(s) \doteq 1$, então

$$\frac{C(s)}{R(s)} \doteq 1$$

Essencialmente, isto significa que a malha interna, usando realimentação positiva, aumentou o ganho direto para um valor muito grande. Conforme estabelecido anteriormente, quando o ganho de malha é muito grande, a função de transferência de malha-fechada $C(s)/R(s)$ torna-se igual ao inverso da função de transferência do elemento de realimentação. Como o sistema indicado na Fig. 5.44(b) possui realimentação unitária, $C(s)/R(s)$ torna-se praticamente igual à unidade. [Portanto, $C(s)/R(s)$ não é sensível a variações de parâmetros de $G(s)$.]

Eliminação de integração. A adição de uma malha menor em volta de um integrador modifica-o para um elemento de retardo de primeira ordem. Considere o sistema indicado na Fig. 5.45(a). A realimentação negativa da saída, conforme a Fig. 5.45(b), modifica o integrador K/s para um elemento de retardo de primeira ordem $K/(s + K)$.

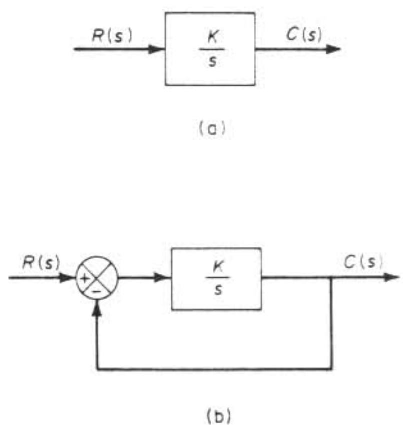


Fig. 5.45 (a) Elemento integrante; (b) elemento de retardo de primeira-ordem.

Comentários sobre o uso de malhas de realimentação. Como foi visto na discussão anterior, o controle por realimentação, o controle de malha-fechada reduz a sensibilidade de um sistema em relação a variações de parâmetros e portanto diminui os efeitos de variações do ganho no ramo direto em resposta a variações da pressão de suprimento, tensão elétrica de alimentação, temperatura etc. No estudo de controladores feito nas Seções 5.2 e 5.3, vimos também que os elementos que

desempenham as várias ações de controle estão na parte de realimentação dos mecanismos controladores e que os elementos de realimentação em um controlador essencialmente aumentam a linearidade do amplificador e aumentam a faixa da sensibilidade proporcional.

O uso de malhas de realimentação nos sistemas de controle, entretanto, aumenta o número de componentes dos sistemas, aumentando a complexidade e introduzindo também a possibilidade de instabilidade.

5.6 TEORIA DE SISTEMAS FLUIDOS

Introdução. Dispositivos fluidos através dos quais fluem ar, gases ou líquidos em canais intrincados e precisos são denominados *dispositivos a fluidos*. Estes dispositivos utilizam componentes sólidos, ou circuitos, para desempenharem as funções de sensores, lógica, amplificação e controle. A teoria de sistemas fluidos, o estudo geral de dispositivos e sistemas a fluidos, é uma das mais novas e mais interessantes áreas desenvolvidas em anos recentes. As vantagens de dispositivos a fluidos são que eles podem ser completamente rudes (devido à ausência de partes móveis mecânicas) e desempenhar funções de controle e cálculo em condições adversas. Dispositivos a fluidos podem operar em amplas faixas de temperatura, temperaturas extremamente altas, sujeitos a choques, vibração, força-G (aceleração) e radiação. (Portanto, dispositivos a fluidos são ideais para operações em locais difíceis e perigosos onde falham outros dispositivos.) Além disso, a simplicidade dos dispositivos básicos a fluidos assegura alta confiabilidade, longa vida e manutenção muito pequena.

Em algumas aplicações de dispositivos a fluidos, o meio controlante é também o meio de fluxo ou potência. Neste caso, o consumo de potência de dispositivos a fluidos não constitui um problema. Nas aplicações onde for limitada a potência disponível, entretanto, como em um sistema de satélite, o consumo relativamente alto de potência de dispositivos a fluido torna-se um problema. Reduzindo-se as dimensões do jato e as pressões de operação, o consumo de potência pode ser reduzido de um certo valor; entretanto, os valores mínimos das dimensões do jato e das pressões de operação são determinados pelos ruídos no sinal e reprodutibilidades de fabricação.

Em relação à velocidade de resposta de dispositivos fluidos, ela é comparável àquelas de dispositivos pneumáticos ou hidráulicos convencionais ou relés eletromecânicos. [Isto significa que a velocidade de resposta de dispositivo a fluidos é muito menor do que aquelas apresentadas por dispositivos eletrônicos. As respostas eletrônicas são geralmente expressas em termo de microssegundos (10^{-6} s) ou nanossegundos (10^{-9} s), enquanto respostas de dispositivos a fluidos são expressas em termo de milissegundos (10^{-3} s).]

Nesta seção discutiremos inicialmente aqueles dispositivos a fluidos denominados amplificadores com aderência de parede, e então apresentaremos a teoria de sistemas digitais a fluidos, uma das mais promissoras aplicações da teoria de sistemas fluidos. Finalmente discutiremos uma aplicação de sistemas digitais a fluidos.

Fenômenos de aderência de parede. O princípio de operação de dispositivos a fluidos é conhecido como interação de jato ou fenômeno de aderência de parede. Um jato fluido pode ser desviado de sua direção normal de fluxo introduzindo um outro jato perpendicular ao primeiro, como indicado na Fig. 5.46. Se o jato de fluido entra em uma câmara relativamente estreita e atinge uma parede, ele "adere" à parede, conforme indicado na Fig. 5.47. Este fenômeno é denominado "efeito de parede". É possível eliminar estes efeitos se aplicarmos um jato ou fluxo na região

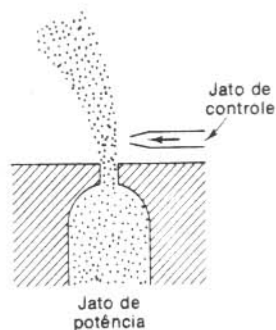


Fig. 5.46 Interação de jato em um dispositivo com fluido.

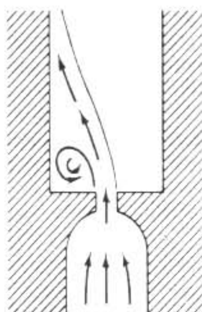


Fig. 5.47 Fenômeno de aderência de parede.

de baixa pressão abaixo do ponto onde o jato atinge a parede. Este fato torna possível projetar um dispositivo biestável, ou *flip-flop*, propiciando jatos de controle em qualquer lado do jato principal, como indicado na Fig. 5.48. (Um dispositivo biestável é aquele que possui duas saídas possíveis e que alternará de uma saída para a outra quando do recebimento de sinais de entrada com fase correta.) Este dispositivo biestável é conveniente para operações lógicas utilizando sinais binários.

Os amplificadores a fluido baseados nos fenômenos de aderência de parede são denominados amplificadores de aderência de parede. Estes amplificadores são normalmente construídos com vidro, plástico, alumínio, latão ou aço inoxidável.

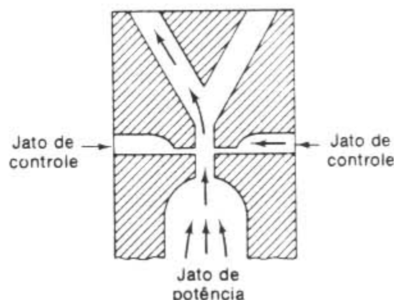


Fig. 5.48 Dispositivo biestável.

Amplificadores biestáveis a fluidos (*flip-flops*). Os amplificadores de aderência de parede são basicamente dispositivos biestáveis. O fluido continuará a fluir em um canal até que seja atuado por um sinal (um jato de fluido) que dirige o fluxo principal de fluido para um outro canal de parede.

A Fig. 5.49 mostra diagramas esquemáticos de um amplificador biestável ou *flip-flop*. Neste elemento, a saída pode ser chaveada de um canal de saída para o

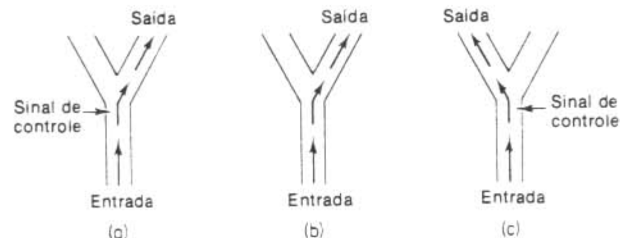


Fig. 5.49 Diagrama esquemático de um amplificador biestável. (a) sinal de controle aplicado à esquerda; (b) sinal de controle removido; (c) sinal de controle aplicado à direita.

outro aplicando-se ar sob pressão no acesso de controle oposto. A saída continuará através deste canal mesmo após ser removido o sinal de controle. O elemento continua a operar como um canal de saída simples até que haja uma variação de sinal. Portanto, este dispositivo possui uma memória. Elementos *flip-flop* servem como banco de memória em um sistema a fluido.

Em dispositivos a fluido, é comumente utilizada para indicar o grau de recuperação uma representação percentual da captura de saída relacionada com a fonte de alimentação, tal como a pressão de saída *versus* a pressão de entrada. Para o elemento *flip-flop* aqui considerado, a máxima pressão de saída é aproximadamente 35% da pressão de suprimento, enquanto o máximo fluxo de saída é aproximadamente 50% do fluxo de suprimento. Os 50% restantes do fluxo de suprimento escoarão através dos compartimentos de exaustão. A pressão mínima de controle que acarreta um chaveamento é aproximadamente 10% da pressão de suprimento.

Amplificadores digitais. Um amplificador digital é similar a um elemento *flip-flop*, exceto que não possui memória. O fluxo de um canal da saída continuará enquanto houver um sinal de controle presente. Um sinal aplicado a qualquer acesso de controle produzirá uma saída no canal de saída oposto.

Operando independentemente de outros elementos com fluidos, o amplificador digital pode ser usado para funções simples, tais como cilindros operantes, relés a ar ou outras unidades lógicas.

Em um amplificador digital, a pressão de controle igual a 10% da pressão de suprimento modificará e manterá o fluxo de saída. A pressão de saída máxima é aproximadamente 50% da pressão de suprimento e o fluxo máximo é aproximadamente 70% do fluxo de suprimento. (O fluxo correspondente a 30% do suprimento restante sairá através dos compartimentos de exaustão.)

Amplificadores proporcionais. Embora os amplificadores de aderência de parede sejam basicamente dispositivos biestáveis, podem ser modificados para dispositivos proporcionais pela ampliação das passagens seguindo o bocal onde ocorre a aderência de parede, conforme indicado na Fig. 5.50. Neste amplificador, o fluxo do jato principal é distribuído entre as duas passagens de saída de acordo com o balanço dos fluxos de controle.

O amplificador proporcional é um dispositivo de saída-entrada diferencial. Referindo-se à Fig. 5.50, notamos que o sinal de entrada é a diferença de pressão existente entre os acessos C_1 e C_2. A saída é a diferença de pressão existente entre os canais O_1 e O_2.

Neste dispositivo o jato é defletido pela diferença de pressão no sinal de entrada que modula a fração do fluxo do jato que cada canal de saída recebe. A Fig.

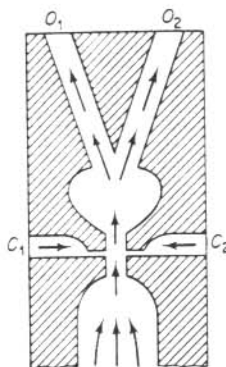


Fig. 5.50 Amplificador proporcional por aderência de parede.

5.51 mostra as curvas características de um amplificador proporcional com aderência de parede.

Para dar uma idéia dos ganhos destes amplificadores, notemos que ganhos de pressão iguais a 10 e ganhos de potência de 100, por estágio, podem ser obtidos. Se dois dispositivos deste tipo são montados em cascata, conforme a Fig. 5.52, então os ganhos são multiplicados.

Como no caso de sistemas elétricos e mecânicos, a oposição total ao fluxo do circuito, representada pela resistência, capacitância e indutância combinadas em uma resultante, é denominada *impedância*. O número de elementos com fluidos que podem ser operados em paralelo de um elemento similar simples é representado pela capacidade de *fan-out*. ("Similar" refere-se à impedância e não aos dispositivos desempenhando a mesma função.) A capacidade de *fan-out* de amplificadores com aderência de parede é aproximadamente quatro.

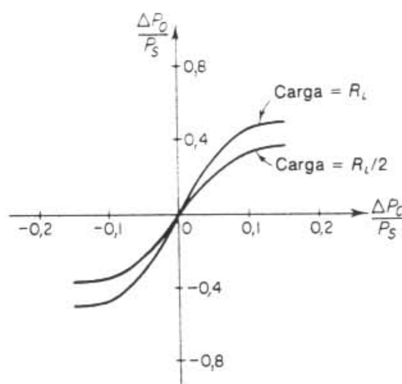


Fig. 5.51 Curvas características de um amplificador proporcional por aderência de parede. (P_s = pressão de suprimento. ΔP_c = diferença de pressão existente entre as portas de controle, ΔP_0 = diferença de pressão existente entre os canais de saída e R_L = resistência de carga.)

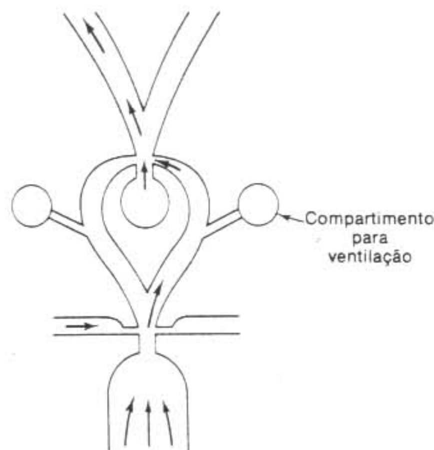


Fig. 5.52 Amplificador em cascata.

Teoria de sistemas digitais fluidos. Dispositivos digitais com fluidos são componentes com fluidos que desempenham funções lógicas, tais como FLIP-FLOP, OR/NOT e AND.

As portas lógicas com fluidos podem ser montadas em circuitos digitais familiares. (Uma porta é um dispositivo ou circuito que permite passagem de um sinal apenas se certos requisitos de controle forem satisfeitos.) A Tabela 5.1 mostra algumas portas lógicas com fluidos e suas respectivas tabelas-verdade. (Uma tabela-verdade é uma correlação tabular de relações entre a entrada e a saída para elementos lógicos.) Como pode ser visto na Tabela 5.1, os dispositivos digitais com fluidos podem permitir, ou inibir, a transmissão de sinal pela aplicação, remoção, ou outras combinações dos sinais de entrada.

A teoria de sistemas digitais fluidos desempenha as mesmas funções lógicas da sua correspondente eletrônica. Há, entretanto, algumas áreas onde cada uma possui vantagens distintas. (Normalmente, a escolha é clara.) Por exemplo, em aplicações onde a confiabilidade em condições extremas (por exemplo, alta temperatura ou radiação) é mais importante que a velocidade de operação, devem ser escolhidos sistemas digitais com fluidos. Se é exigida operação lógica com alta velocidade em condições normais, são preferíveis os dispositivos eletrônicos. Aplicações de sistemas digitais com fluidos são quase sempre encontradas em armazenagem, alimentação de máquinas, seqüenciadoras, operadoras etc.

Muitas funções de relés de controle podem ser desempenhadas por meio de dispositivos com fluidos, e muitas funções lógicas de relés e sistemas eletrônicos podem ser implementadas através de dispositivos com fluido.

Uma aplicação típica de sistema com fluido*. Componentes com fluidos podem sensorar posições, propiciar comandos de operações, desempenhar engrenamentos lógicos e controlar, por meio de válvulas, o sistema de controle de máquinas.

Considere o sistema mecânico indicado na Fig. 5.53. Deseja-se projetar um sistema lógico de controle para manter a seqüência de uma furadeira, de modo que, quando a broca está retraída e um botão é atuado, a furadeira avança para baixo. A broca avança descendentemente e então retorna à sua posição para cima a fim de

*Este exemplo é obtido de *Fluidics-Plugging the Control Capability Gap* por The General Purpose Control Department, General Electric Company, Bloomington, Illinois.

Tabela 5.1 Portas lógicas com fluido e suas respectivas tabelas-verdade

Tabela-verdade	Dispositivo com fluido	Símbolo
	FLIP-FLOP	
	OR/NOT	
	AND	

FLIP-FLOP

Estado presente				Próximo estado	
C_1	C_2	O_1	O_2	O_1	O_2
0	0	0	1	0	1
0	0	1	0	1	0
0	1	0	1	1	0
0	1	1	0	1	0
1	0	0	1	0	1
1	0	1	0	0	1

OR/NOT

C_1	C_2	C_3	O_1	O_2
0	0	0	1	0
Todos os outros estados			0	1

AND

C_1	C_2	O_1	O_2
0	0	0	1
0	1	0	1
1	0	0	1
1	1	1	0

completar o ciclo. Para retorno de emergência, exige-se um botão para retrair a broca em qualquer instante.

Apresentaremos uma solução para este problema utilizando um sistema digital com fluidos. A Fig. 5.54 mostra um circuito lógico que pode desempenhar a seqüência de operações desejada. O elemento FLIP-FLOP 1 é acionado pelo botão de início (partida). O botão de início é soltado, porém o elemento continuará a fornecer a saída desejada. A saída do elemento FLIP-FLOP 1 alimenta o elemento AND, que combina este sinal com o outro sinal proveniente da chave de limite

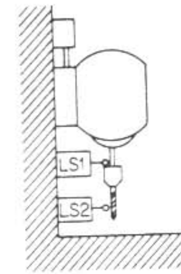


Fig. 5.53 Sistema mecânico.

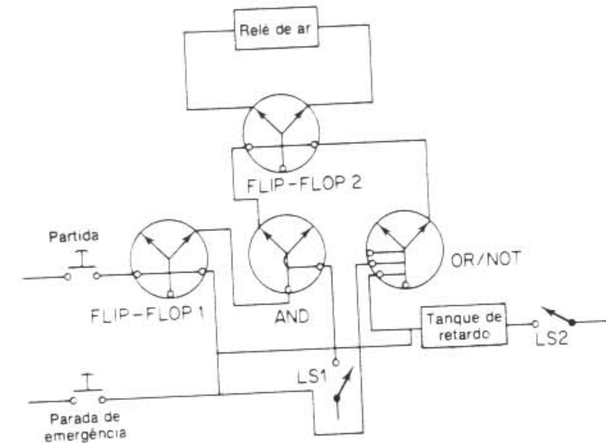


Fig. 5.54 Circuito lógico.

(LS 1). Este sinal ativa o FLIP-FLOP 2, que desloca o relé de ar e inicia o movimento da broca para baixo. O suporte de broca move-se para baixo e opera a broca até que o sinal proveniente de LS 2 atinja o tanque de retardo. Este desloca o elemento OR/NOT, que por sua vez chaveia o FLIP-FLOP 2 que aciona o relé de ar para a sua posição original. O suporte de broca move-se agora para cima. O ciclo repetir-se-á quando o botão de início for empurrado e LS 1 for ativado.

Se o botão de parada de emergência for acionado em qualquer instante, ele acionará o elemento OR/NOT e o FLIP-FLOP 1. O elemento OR/NOT acionará o FLIP-FLOP 2, o que causará a imediata retração do suporte da broca.

PROBLEMAS ILUSTRATIVOS E SOLUÇÕES

Problema A.5.1 O termo normalmente usado para definir o ganho ou sensibilidade de um controlador proporcional é a *banda proporcional*. Esta é a variação percentual na entrada para o controlador (sinal erro) necessária para causar 100% de variação na saída do atuador.

Portanto, uma banda proporcional pequena corresponde a um alto ganho ou alta sensibilidade proporcional.

Qual é a banda proporcional se o controlador e atuador tiverem um ganho global de 4%/%? (Note que as variações totais na entrada do controlador e na saída do atuador são dadas como 100%. Portanto um ganho de 4%/% significa que há uma variação de 4% na saída se a variação na entrada for 1%.)

Solução

$$\text{Banda proporcional} = \frac{100\%}{\text{ganho em }\%/\%} = \frac{100\%}{4\%/\%} = 25\%$$

Problema A.5.2 Considere o sistema de controle de nível de líquido indicado na Fig. 5.55. Suponha que o ponto de ajuste do controlador está fixado. Supondo um distúrbio em degrau de valor n_0, determine o erro. Suponha que n_0 seja pequeno e as variações nas variáveis em relação a seus valores de regime estacionário respectivos também sejam pequenas. O controlador é do tipo proporcional.

Se o controlador não for proporcional, mas integral, qual será o erro em regime estacionário?

Solução. A Fig. 5.56 representa um diagrama de blocos do sistema quando o controlador é proporcional com ganho K_p. (Suponha que a função de transferência da válvula pneumática

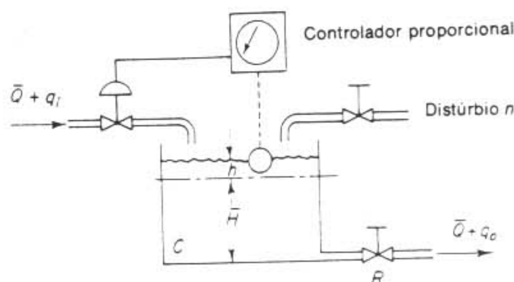


Fig. 5.55 Sistema de controle de nível de líquido.

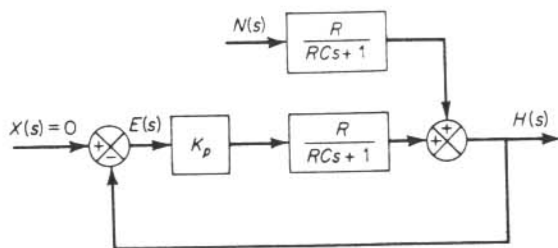


Fig. 5.56 Diagrama de blocos do sistema de controle de nível de líquido indicado na Fig. 5.55.

seja unitária. Desde que o ponto de ajuste está fixado, a variação no ponto de ajuste é nula, ou $X(s) = 0$. A transformada de Laplace de $h(t)$ é

$$H(s) = \frac{K_pR}{RCs+1}E(s) + \frac{R}{RCs+1}N(s)$$

Então,

$$E(s) = -H(s) = -\frac{K_pR}{RCs+1}E(s) - \frac{R}{RCs+1}N(s)$$

Portanto,

$$E(s) = -\frac{R}{RCs+1+K_pR}N(s)$$

Desde que,

$$N(s) = \frac{n_0}{s}$$

obtemos

$$\begin{aligned} E(s) &= -\frac{R}{RCs+1+K_pR}\frac{n_0}{s} \\ &= \frac{Rn_0}{1-K_pR}\left(\frac{1}{s+\dfrac{1+K_pR}{RC}}\right) - \frac{Rn_0}{1+K_pR}\frac{1}{s} \end{aligned}$$

A solução temporal para $t > 0$ é

$$e(t) = \frac{Rn_0}{1-K_pR}\left[\exp\left(-\frac{1+K_pR}{RC}t\right) - 1\right]$$

Portanto, a constante de tempo é $RC/(1+K_pR)$. (Na ausência do controlador, a constante de tempo é igual a RC.) Conforme aumenta o ganho do controlador, diminui a constante de tempo. O erro em regime estacionário é

$$e(\infty) = -\frac{Rn_0}{1+K_pR}$$

Conforme o ganho K_p do controlador é aumentado, o erro em regime estacionário, ou o desajuste, é reduzido. Portanto, matematicamente, quanto maior for o ganho K_p, menores serão o desajuste e a constante de tempo. Em sistemas práticos, entretanto, se o ganho K_p do controlador proporcional for aumentado para um valor muito grande, deve resultar uma oscilação na saída desde que desprezamos em nossa análise todos os pequenos atrasos e pequenas constantes de tempo que podem existir no sistema de controle real. (Se estes pequenos atrasos e constantes de tempo forem incluídos na análise, a função de transferência torna-se de maior ordem e, para valores muito grandes de K_p, há a possibilidade de ocorrer oscilação ou mesmo instabilidade.)

Se o controlador for integral, supondo que a função de transferência do controlador seja

$$G_c = \frac{K}{s}$$

obtemos

$$E(s) = -\frac{Rs}{RCs^2 + s + KR}N(s)$$

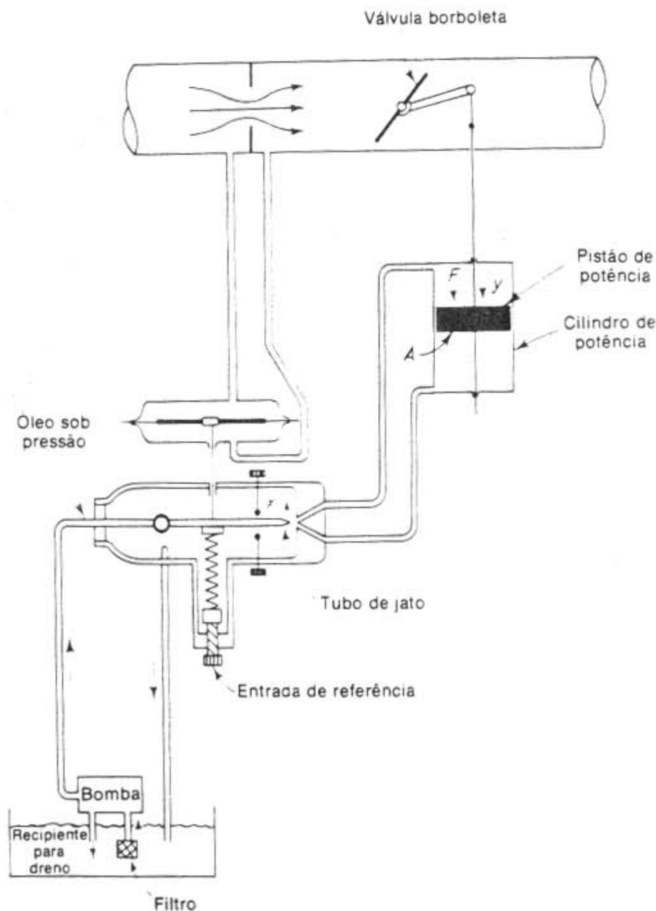


Fig. 5.57 Diagrama esquemático de um sistema de controle de fluxo usando um controlador de tubo de jato hidráulico.

O erro em regime estacionário para um distúrbio em degrau $N(s) = n_0/s$ é

$$\begin{aligned} e(\infty) &= \lim_{s \to 0} sE(s) \\ &= \lim_{s \to 0} \frac{-Rs^2}{RCs^2 + s + KR}\frac{n_0}{s} \\ &= 0 \end{aligned}$$

Portanto, um controlador integral elimina o erro em regime estacionário, ou desajuste, devido ao distúrbio em degrau. (O valor de K deve ser escolhido de modo que a resposta transitória devida à entrada de comando e/ou distúrbio amorteçam com uma velocidade razoável. Vide Cap. 6 para análise de resposta transitória.)

Problema A.5.3 A Fig. 5.57 mostra um controlador de cano de jato hidráulico aplicado a um sistema de controle de fluxo. O controlador de cano de jato governa a posição da válvula borboleta. Discuta a operação deste sistema. Construa um gráfico da curva possível relacionando o deslocamento x do bocal em relação à força total F atuando no pistão de potência.

Solução. A operação deste sistema é a seguinte: A taxa de fluxo é medida pelo orifício, e a diferença de pressão produzida por este orifício é transmitida ao diafragma do dispositivo medidor de pressão. O diafragma é ligado ao bocal de giro livre, ou cano de jato, através de uma ligação. O óleo a alta pressão é ejetado do bocal durante todo o tempo. Quando o bocal está na posição neutra, não flui óleo através de qualquer um dos canos para mover o pistão de potência. Se o bocal é deslocado pelo movimento do braço de balanço para um lado, o óleo sob alta pressão flui através do cano correspondente, e o óleo no cilindro de potência flui para o reservatório de dreno através do outro cano.

Suponha que o sistema esteja inicialmente em repouso. Se a entrada de referência for variada subitamente para uma taxa de fluxo maior, então o bocal é movido de modo tal a mover o pistão de potência e abrir a válvula borboleta. Haverá então um aumento na taxa de fluxo, a diferença de pressão através do orifício torna-se maior e o bocal retorna à posição neutra. O movimento do pistão de potência cessa quando x, o deslocamento do bocal, retorna e pára na posição neutra. (O controlador do cano de jato, portanto, possui uma propriedade de integração.)

A relação entre a força total F agindo no pistão de potência e o deslocamento x do bocal é indicada na Fig. 5.58. A força total é igual à diferença de pressão ΔP através do pistão vezes a área A do pistão de potência. Para um pequeno deslocamento x do bocal, a força total F e o deslocamento x podem ser considerados proporcionais.

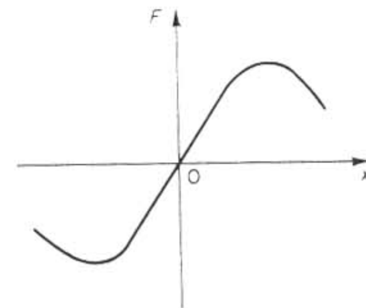


Fig. 5.58 Curva força *versus* deslocamento.

Problema A.5.4 Desenhe um diagrama de blocos do controlador pneumático indicado na Fig. 5.59. Deduza então a função de transferência deste controlador.

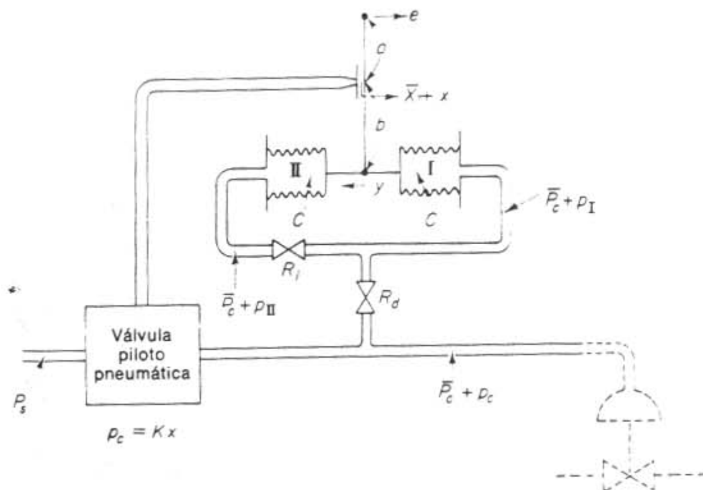


Fig. 5.59 Diagrama esquemático de um controlador pneumático.

Se a resistência R_d for removida (substituída pela tubulação de linha dimensionada), que ação de controle resultará? Se a resistência R_i for removida (substituída pela tubulação de linha dimensionada), que ação de controle resultará?

Solução. Vamos supor que quando $e = 0$, a distância haste-bocal é igual a $\bar{X}$ e a pressão de controle é igual a $\bar{P}_c$. Na presente análise, suporemos pequenos desvios em relação aos respectivos valores de referência como segue:

e = pequeno sinal erro
x = pequena variação na distância bocal-haste
p_c = pequena variação na pressão de controle
p_I = pequena variação de pressão no fole I devida a pequena variação na pressão de controle
p_{II} = pequena variação de pressão no fole II devida a pequena variação na pressão de controle
y = pequeno deslocamento na extremidade inferior da haste

Neste controlador, p_c é transmitida ao fole I através da resistência R_d. Analogamente, p_c é transmitida ao fole II através da série de resistências R_d e R_i. Uma relação aproximada entre p_I e p_c é

$$\frac{P_I(s)}{P_c(s)} = \frac{1}{R_dCs + 1} = \frac{1}{T_ds + 1}$$

onde

$$T_d = R_dC = \text{tempo derivativo}$$

Analogamente, p_{II} e p_I são relacionadas pela função de transferência

$$\frac{P_{II}(s)}{P_I(s)} = \frac{1}{R_iCs + 1} = \frac{1}{T_is + 1}$$

onde

$$T_i = R_iC = \text{tempo integral}$$

A equação de equilíbrio de forças para os dois foles é

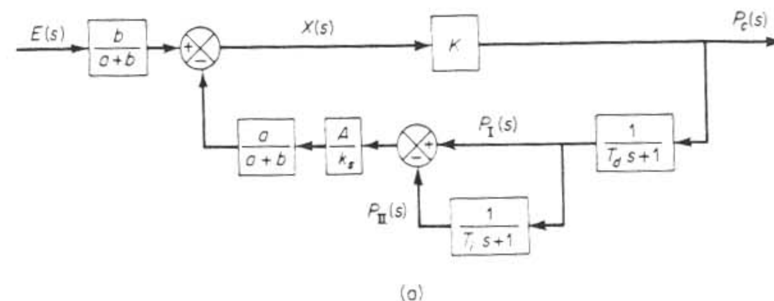
$$(p_I - p_{II})A = k_sy$$

onde k_s é a dureza dos dois foles interligados e A é a área da seção transversal dos foles. A relação entre as variáveis e, x e y é

$$x = \frac{b}{a+b}e - \frac{a}{a+b}y$$

A relação entre p_c e x é

$$p_c = Kx$$



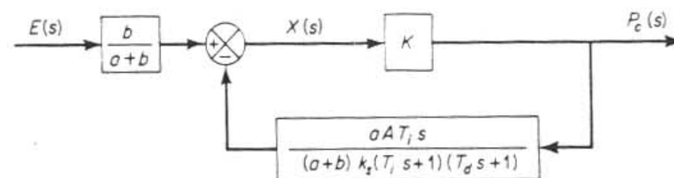


Fig. 5.60 (a) Diagrama de blocos do controlador pneumático mostrado na Fig. 5.59; (b) diagrama de blocos simplificado.

Das equações obtidas, pode ser desenhado um diagrama de blocos do controlador conforme indicado na Fig. 5.60(a). A simplificação deste diagrama de blocos resulta na Fig. 5.60(b).

A função de transferência entre $P_c(s)$ e $E(s)$ é

$$\frac{P_c(s)}{E(s)} = \frac{\dfrac{b}{a-b}K}{1 - K\dfrac{a}{a+b}\dfrac{A}{k_s}\left(\dfrac{T_i s}{T_i s + 1}\right)\left(\dfrac{1}{T_d s + 1}\right)}$$

Para um controlador prático, sob operação normal $|KaAT_is/[(a+b)k_s(T_is+1)(T_ds+1)]|$ é muito maior do que a unidade e $T_i >> T_d$. Portanto, a função de transferência pode ser simplificada como segue:

$$\frac{P_c(s)}{E(s)} \doteq \frac{bk_s(T_i s + 1)(T_d s + 1)}{aAT_i s}$$

$$= \frac{bk_s}{aA}\left(\frac{T_i + T_d}{T_i} + T_d s + \frac{1}{T_i s}\right)$$

$$\doteq K_p\left(1 + T_d s + \frac{1}{T_i s}\right)$$

onde

$$K_p = \frac{bk_s}{aA}$$

Portanto o controlador indicado na Fig. 5.59 é proporcional-mais-derivativo-mais-integral.

Se a resistência R_d for removida, ou $R_d = 0$, a ação resultante é a de um controlador proporcional-mais-integral.

Se a resistência R_i for removida, ou $R_i = 0$, a ação que resulta é a de um controlador proporcional de banda estreita ou de duas posições. (Note que as ações dos dois foles de realimentação se cancelam e não há realimentação.)

Problema A.5.5 Obtenha a função de transferência do controlador indicado na Fig. 5.33(c).

Solução. A Fig. 5.61 mostra o circuito de realimentação. As equações para este circuito são

$$\frac{1}{C_d s}[I_1(s) - I_2(s)] + R_d I_1(s) = E_0(s)$$

$$\frac{1}{C_d s}[I_2(s) - I_1(s)] + \frac{1}{C_i s}I_2(s) + R_i I_2(s) = 0$$

Portanto,

$$\frac{I_2(s)}{E_0(s)} = \frac{C_i s}{R_i C_i R_d C_d s^2 + (R_i C_i + R_d C_i + R_d C_d)s + 1}$$

ou

$$\frac{E_1(s)}{E_0(s)} = \frac{R_i C_i s}{R_i C_i R_d C_d s^2 + (R_i C_i + R_d C_i + R_d C_d)s + 1}$$

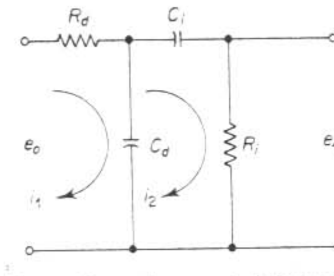


Fig. 5.61 Circuito de realimentação usado no controlador mostrado na Fig. 5.33(c).

Verificando que na Fig. 5.33(c)

$$(e_i - e_f)K = e_0, \qquad e_f = e_1\frac{R_2}{R_1}$$

obtemos

$$\left[E_i(s) - \frac{R_2}{R_1}\left(\frac{R_i C_i s E_0(s)}{R_i C_i R_d C_d s^2 + (R_i C_i + R_d C_i + R_d C_d)s + 1}\right)\right]K = E_0(s)$$

A função de transferência $E_0(s)/E_i(s)$ é

$$\frac{E_0(s)}{E_i(s)} = \frac{KR_1[R_i C_i R_d C_d s^2 + (R_i C_i + R_d C_i + R_d C_d)s + 1]}{KR_2 R_i C_i s + R_1[R_i C_i R_d C_d s^2 + (R_i C_i + R_d C_i + R_d C_d)s + 1]}$$

Se o ganho de malha é muito maior do que a unidade, então esta última equação pode ser simplificada para

$$\frac{E_0(s)}{E_i(s)} = \frac{R_1[R_i C_i R_d C_d s^2 + (R_i C_i + R_d C_i + R_d C_d)s + 1]}{R_2 R_i C_i s}$$

$$= K_p\left[T_d s + \left(1 + \frac{R_d}{R_i} + \frac{T_d}{T_i}\right) + \frac{1}{T_i s}\right]$$

onde

$$K_p = R_1/R_2, \qquad T_d = R_d C_d, \qquad T_i = R_i C_i$$

Definindo

$$\alpha = 1 + \frac{R_d}{R_i} + \frac{T_d}{T_i}$$

Então,

$$\frac{E_0(s)}{E_i(s)} = K_p\alpha\left(1 + \frac{T_d}{\alpha}s + \frac{1}{\alpha T_i s}\right)$$

Portanto, o controlador é proporcional-mais-derivativo-mais-integral.

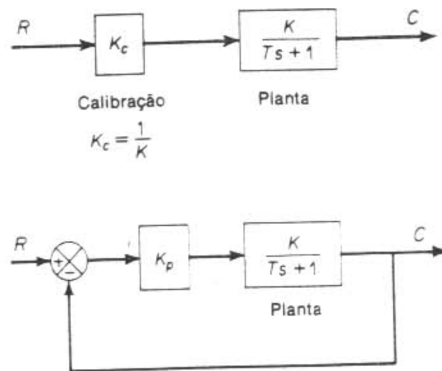


Fig. 5.62 Diagramas de blocos de um sistema de controle de malha-aberta e um sistema de controle de malha-fechada.

Problema A.5.6 Considere o sistema de controle de malha-aberta e o sistema de controle de malha-fechada indicados na Fig. 5.62. No sistema de malha-aberta o ganho K_c é calibrado de modo que $K_c = 1/K$. Portanto, a função de transferência do sistema de controle de malha-aberta é

$$G_0(s) = \frac{1}{K}\frac{K}{Ts-1} = \frac{1}{Ts+1}$$

No sistema de controle de malha-fechada, o ganho K_p do controlador é ajustado de modo que $K_pK >> 1$.

Supondo uma entrada com degrau unitário, compare os erros em regime estacionário para estes sistemas de controle.

Solução. Para o sistema de controle de malha-aberta, o sinal erro é

$$e(t) = r(t) - c(t)$$

ou

$$\begin{aligned} E(s) &= R(s) - C(s) \\ &= [1 - G_0(s)]R(s) \end{aligned}$$

O erro de regime estacionário para a resposta de entrada ao degrau é

$$\begin{aligned} e_{ss} &= \lim_{s\to 0} sE(s) \\ &= \lim_{s\to 0} s[1 - G_0(s)]\frac{1}{s} \\ &= 1 - G_0(0) \end{aligned}$$

Se $G_0(0)$, o ganho c.c. do sistema de controle de malha aberta, é igual à unidade, então o erro em regime estacionário é nulo. Devido a variações ambientais e envelhecimento de componentes, entretanto, o ganho c.c. $G_0(0)$ apresentará deriva em relação ao zero conforme

transcorre o tempo e o erro em regime estacionário não permanecerá nulo por muito tempo. Este erro em regime estacionário em um sistema de controle de malha-aberta persistirá até que o sistema seja recalibrado.

Para o sistema de controle de malha-fechada, o sinal erro é

$$\begin{aligned} E(s) &= R(s) - C(s) \\ &= \frac{1}{1+G(s)}R(s) \end{aligned}$$

onde

$$G(s) = \frac{K_pK}{Ts+1}$$

O erro de regime estacionário para uma entrada em degrau é

$$\begin{aligned} e_{ss} &= \lim_{s\to 0} s\left[\frac{1}{1+G(s)}\right]\frac{1}{s} \\ &= \frac{1}{1+G(0)} \\ &= \frac{1}{1+K_pK} \end{aligned}$$

No sistema de controle de malha fechada, o ganho K_p é ajustado em um valor muito maior do que $1/K$. Portanto o erro em regime estacionário pode ser feito pequeno, embora não nulo.

Vamos supor a seguinte variação na função de transferência da planta, supondo K_c e K_p constantes:

$$\frac{K+\Delta K}{Ts+1}$$

Por simplicidade, vamos supor $K = 10$, $\Delta K = 1$, ou $\Delta K/K = 0,1$. Então o erro em regime estacionário para uma entrada em degrau unitário no sistema de controle de malha-aberta resulta

$$\begin{aligned} e_{ss} &= 1 - \frac{1}{K}(K+\Delta K) \\ &= 1 - 1,1 \\ &= -0,1 \end{aligned}$$

No sistema de controle de malha-fechada, se K_p é ajustado para $100/K$, então o erro em regime estacionário para uma entrada em degrau unitário resulta

$$\begin{aligned} e_{ss} &= \frac{1}{1+G(0)} \\ &= \frac{1}{1+\frac{100}{K}(K+\Delta K)} \\ &= \frac{1}{1+110} \\ &= 0,009 \end{aligned}$$

Portanto, o sistema de controle de malha-fechada é superior ao sistema de controle de malha-aberta na presença de variações ambientais, envelhecimento de componentes etc. que, definitivamente, afetam o desempenho em regime estacionário.

Problema A.5.7 O diagrama de blocos da Fig. 5.63 mostra um sistema de controle de velocidade no qual o elemento de saída do sistema é sujeito a um distúrbio de torque. No diagrama, $\Omega_r(s)$, $\Omega(s)$, $T(s)$ e $N(s)$ correspondem às transformadas de Laplace da velocidade de referência, velocidade de saída, torque de excitação e torque de distúrbio, respectivamente. Na ausência de um torque de distúrbio, a velocidade de saída é igual à velocidade de referência.

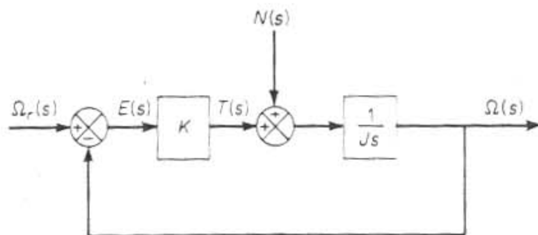


Fig. 5.63 Diagrama de blocos de um sistema de controle de velocidade.

Investigue a resposta deste sistema ao torque de distúrbio de degrau unitário. Suponha que a entrada de referência é nula, ou $\Omega_r(s) = 0$.

Solução. A Fig. 5.64 apresenta um diagrama de blocos modificado e conveniente para a presente análise. A função de transferência de malha-fechada é

$$\frac{\Omega_N(s)}{N(s)} = \frac{1}{Js + K}$$

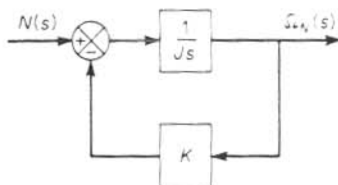


Fig. 5.64 Diagrama de blocos do sistema de controle de velocidade da Fig. 5.63 quando $\Omega_r(s) = 0$.

onde $\Omega_N(s)$ é a transformada de Laplace da velocidade de saída devida ao torque de distúrbio. Para um torque de distúrbio com entrada unitária, a velocidade de saída em regime estacionário é

$$\begin{aligned}\omega_N(\infty) &= \lim_{s\to 0} s\Omega_N(s) \\ &= \lim_{s\to 0} \frac{s}{Js + K}\frac{1}{s} \\ &= \frac{1}{K}\end{aligned}$$

Desta análise concluímos que, se um torque de distúrbio em degrau for aplicado ao elemento de saída do sistema, resultará uma velocidade de erro, de modo que o torque motor cancelará exatamente o torque do distúrbio. Para desenvolver este torque motor é necessário que haja um erro na velocidade de modo a resultar um torque não nulo.

Problema A.5.8 No sistema considerado no Problema A.5.7, deseja-se eliminar tanto quanto possível os erros de velocidade devidos a distúrbios de torques.

É possível cancelar o efeito de um distúrbio de torque em regime estacionário de modo que um distúrbio de torque constante aplicado ao elemento de saída não acarrete variação de velocidade em regime estacionário?

Solução. Suponha que escolhemos um controlador conveniente cuja função de transferência é $G_c(s)$, como indicado na Fig. 5.65. Então, na ausência da entrada de referência, a função de

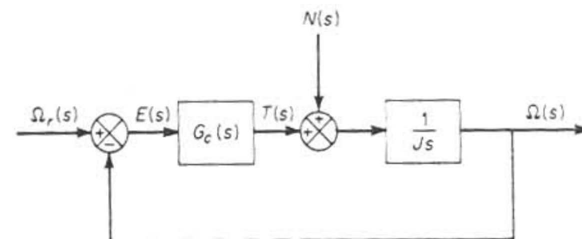


Fig. 5.65 Diagrama de blocos de um sistema de controle de velocidade.

transferência entre a velocidade de saída $\Omega_N(s)$ e o torque de distúrbio $N(s)$ é

$$\begin{aligned}\frac{\Omega_N(s)}{N(s)} &= \frac{\frac{1}{Js}}{1 + \frac{1}{Js}G_c(s)} \\ &= \frac{1}{Js + G_c(s)}\end{aligned}$$

A velocidade de saída em regime estacionário devida a um torque de distúrbio em degrau unitário é

$$\begin{aligned}\omega_N(\infty) &= \lim_{s\to 0} s\Omega_N(s) \\ &= \lim_{s\to 0} \frac{s}{Js + G_c(s)}\frac{1}{s} \\ &= \frac{1}{G_c(0)}\end{aligned}$$

De maneira a satisfazer o requisito de

$$\omega_N(\infty) = 0$$

devemos escolher $G_c(0) = \infty$. Isto pode ser realizado se escolhermos

$$G_c(s) = \frac{K}{s}$$

A ação de controle integral continuará a corrigir até que o erro seja nulo. Este controlador, entretanto, apresenta um problema de estabilidade porque a equação característica possui duas raízes imaginárias.

Um método para estabilizar este sistema é adicionar um modo proporcional ao controlador, ou escolher

$$G_c(s) = K_p + \frac{K}{s}$$

Com este controlador, o diagrama de blocos da Fig. 5.65 na ausência da entrada de referência pode ser modificado para o da Fig. 5.66. A função de transferência de malha-fechada

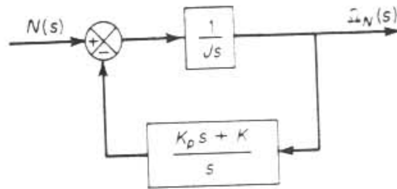


Fig. 5.66 Diagrama de blocos do sistema de controle de velocidade da Fig. 5.65 quando $G_c(s) = K_p + (K/s)$ e $\Omega_r(s) = 0$.

$\Omega_N(s)/N(s)$ resulta

$$\frac{\Omega_N(s)}{N(s)} = \frac{s}{Js^2 + K_p s + K}$$

Para um torque de distúrbio com degrau unitário, a velocidade de saída em regime estacionário é

$$\begin{aligned} \omega_N(\infty) &= \lim_{s\to 0} s\Omega_N(s) \\ &= \lim_{s\to 0} \frac{s^2}{Js^2 + K_p s + K}\frac{1}{s} \\ &= 0 \end{aligned}$$

Portanto, verificamos que o controlador proporcional-mais-integral elimina o erro de velocidade em regime estacionário.

O uso da ação de controle integral aumentou a ordem do sistema de uma unidade. (Isto tende a produzir uma resposta oscilatória.)

No sistema presente, um torque de distúrbio em degrau acarretará um erro transitório na velocidade de saída, porém o erro resulta nulo em regime estacionário. O integrador fornece uma saída não nula com erro nulo. (A saída não nula do integrador produz um torque motor que cancela exatamente o torque do distúrbio.)

Note que o integrador na função de transferência da planta não elimina o erro em regime estacionário devido ao torque do distúrbio. Para eliminar isto, deve haver um integrador antes do ponto onde entra o torque do distúrbio.

Problema A.5.9 A Fig. 5.67(a) apresenta um diagrama esquemático de um dispositivo de controle pneumático. Consiste em quatro câmaras separadas por três diafragmas que são rigidamente ligados conforme indicado no diagrama. Este dispositivo possui dois tubos de entrada onde são conectadas duas pressões de controle P_1 e P_2. Se tais pressões forem aplicadas, a montagem de diafragmas distorce e interrompe uma das duas entradas de fornecimento. A pressão de saída P_0 é então igual ou a P_s ou a 0. A pressão de saída P_0 pode ser indicada em um gráfico *versus* a diferença de pressão $P_1 - P_2$, conforme a Fig. 5.67(b). Este dispositivo pode ser utilizado como um dispositivo lógico.

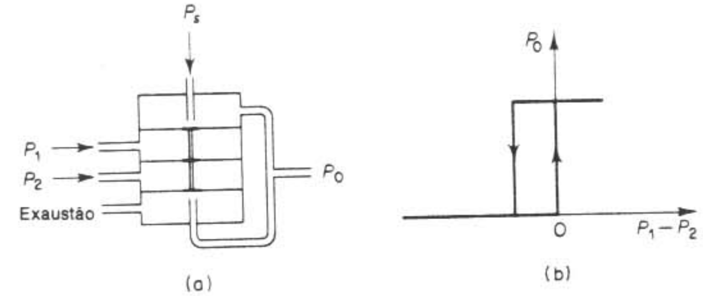


Fig. 5.67 (a) Diagrama esquemático de um dispositivo de controle pneumático; (b) curva característica.

Se uma pressão de polarização for aplicada a uma das câmaras, conforme a Fig. 5.68(a), a curva característica é deslocada conforme a Fig. 5.68(b).

A Fig. 5.69 mostra uma combinação dos dois dispositivos. Que operação lógica pode ser desempenhada por esta combinação? Supondo que a pressão de suprimento é igual a 1 e as pressões de entrada nos pontos *A* e *B* sejam ou 0 ou 1, determine a pressão de saída no ponto *F*.

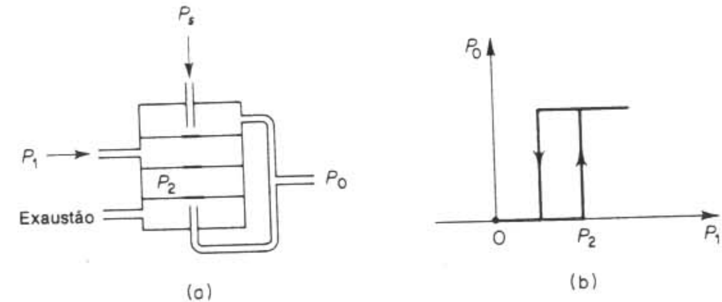


Fig. 5.68 (a) Diagrama esquemático de um dispositivo de controle pneumático; (b) curva característica.

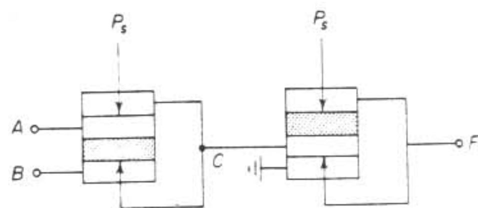


Fig. 5.69 Combinação de dois dispositivos de controle pneumáticos.

Solução. Da Fig. 5.69, obtemos a seguinte tabela:

A	B	C	F
0	0	0	1
1	0	1	0
0	1	1	0
1	1	1	0

Portanto, a operação lógica obtida é a NOR: isto é, $F = \bar{A}.\bar{B}$.

PROBLEMAS

Problema B.5.1 O diagrama esquemático de um controlador pneumático é indicado na Fig. 5.70. Desenhe um diagrama de blocos do controlador e posteriormente deduza a função de transferência.

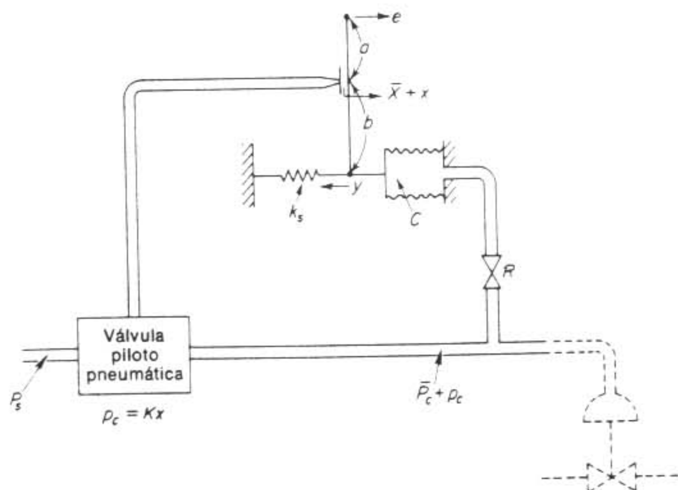


Fig. 5.70 Diagrama esquemático de um controlador pneumático.

Problema B.5.2 A Fig. 5.71 mostra o diagrama esquemático de um servomotor hidráulico no qual o sinal erro é amplificado em dois estágios. Desenhe um diagrama de blocos do sistema e determine, posteriormente, a função de transferência entre y e x.

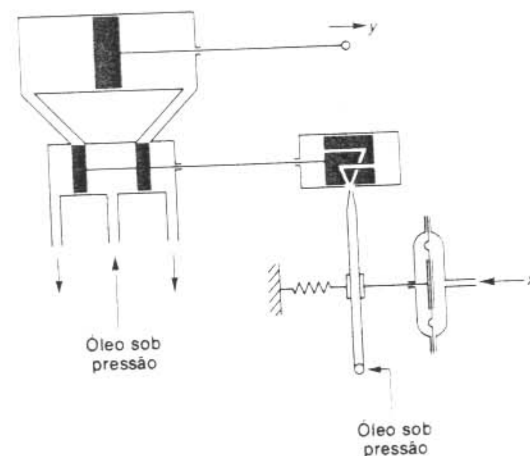


Fig. 5.71 Diagrama esquemático de um servomotor hidráulico.

Problema B.5.3 A Fig. 5.72 fornece o diagrama esquemático de um servomotor hidráulico. Que tipo de ação de controle produz este servomotor?

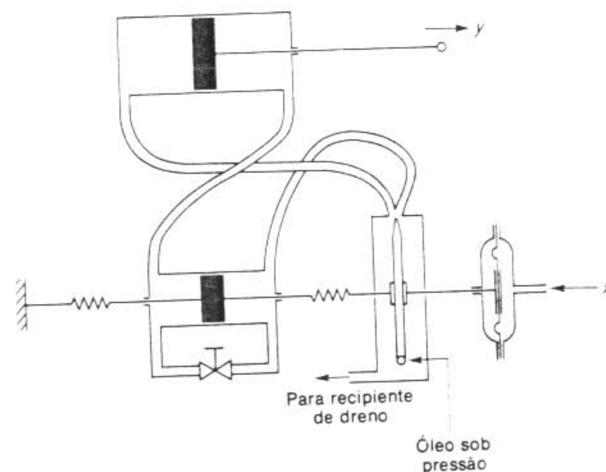


Fig. 5.72 Diagrama esquemático de um servomotor hidráulico.

Problema B.5.4 A Fig. 5.73 mostra o diagrama esquemático de um sistema de controle de velocidade. O eixo é excitado a uma velocidade ω através de uma redução por engrenagens. Os volantes girantes resultam em uma força centrífuga que é oposta à força da mola. A velocidade desejada é ajustada pela mola de pré-carga. Que tipo de ação de controle produz o controlador?

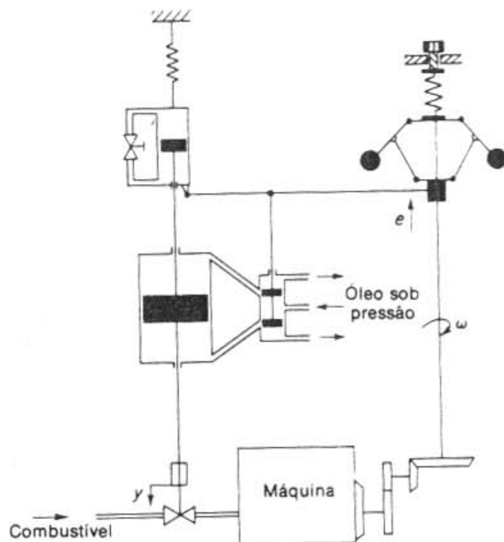


Fig. 5.73 Diagrama esquemático de um sistema de controle de velocidade.

Problema B.5.5 Considere o controlador indicado na Fig. 5.74. Que tipo de ação de controle exerce este controlador?

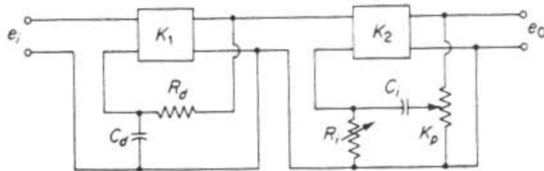
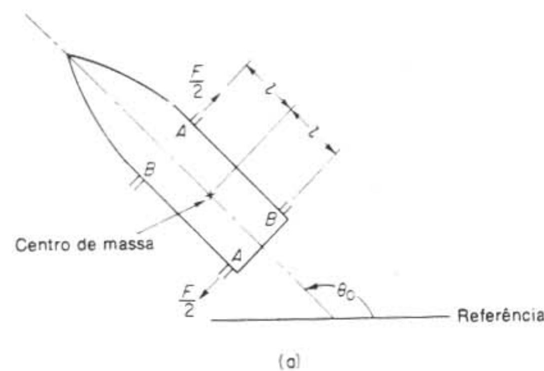


Fig. 5.74 Diagrama esquemático de um controlador eletrônico.

Problema B.5.6 A Fig. 5.75(a) apresenta o diagrama esquemático de um sistema de controle de atitude de satélite. Pequenos jatos aplicam força de reação para girar o corpo do satélite na atitude desejada. Os dois jatos posicionados anti-simetricamente e indicados por A e B operam em pares. Suponha que cada jato aplique ao sistema uma força $F/2$ e um torque Fl. O momento de inércia em relação ao centro de massa é J.

Suponha que o controlador de atitude seja do tipo proporcional-mais-derivativo. A representação do sistema em diagrama de blocos é fornecida na Fig. 5.75(b). Determine o valor do tempo derivativo de modo que a relação de amortecimento ζ seja 0.7.



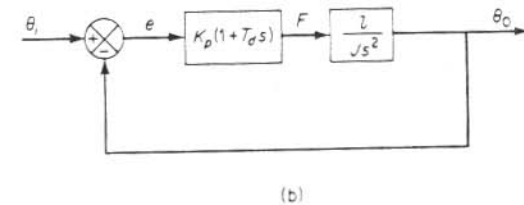


Fig. 5.75 (a) Diagrama esquemático de um sistema de controle de atitude de satélite; (b) diagrama de bloco.

Problema B.5.7 Considere o sistema de múltiplos laços indicado na Fig. 5.76. Obtenha a função de transferência de malha-fechada entre $C(s)$ e $N(s)$.

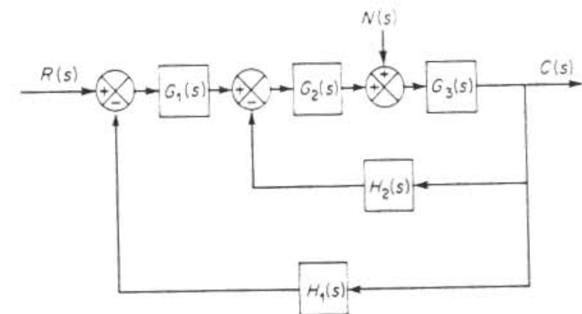


Fig. 5.76 Sistema de múltiplas malhas.

Problema B.5.8 Se o caminho direto de um sistema de controle contém pelo menos um elemento integrante, então a saída continua a variar enquanto existir um erro. A saída pára quando o erro é precisamente zero. Se um distúrbio externo entra no sistema, é desejável ter um elemento integrante entre o elemento de medida do erro e o ponto onde entra o distúrbio, de modo que o efeito do distúrbio externo seja feito nulo em regime estacionário.

Mostre que se o distúrbio é uma função rampa, então o erro em regime estacionário, devido a este distúrbio em rampa, pode ser eliminado somente se dois integradores precederem o ponto onde entra o distúrbio.

Problema B.5.9 Considere o sistema indicado na Fig. 5.77(a) onde K é um ganho ajustável e $G(s)$ e $H(s)$ são componentes fixos. A função de transferência de malha-fechada para o distúrbio é

$$\frac{C(s)}{N(s)} = \frac{1}{1 + KG(s)H(s)}$$

Para minimizar o efeito dos distúrbios, o ganho ajustável K deve ser escolhido tão grande quanto possível.

Isto é também verdade para o sistema mostrado na Fig. 5.77(b)?

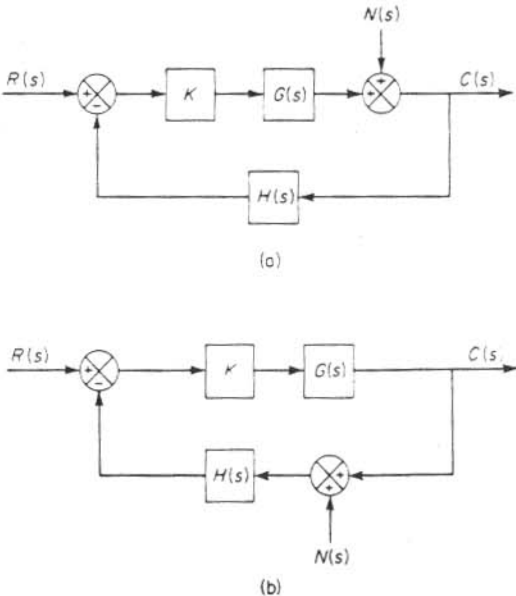


Fig. 5.77 (a) Sistema de controle com distúrbio entrando no ramo direto; (b) sistema de controle com distúrbio entrando no ramo de realimentação.

Problema B.5.10 A Fig. 5.78 mostra o diagrama esquemático de um dispositivo com fluido. Que função este dispositivo desempenha?

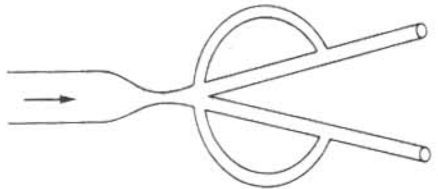


Fig. 5.78 Diagrama esquemático de um dispositivo com fluido.

Problema B.5.11 Um resistor com fluido é um elemento passivo com fluido que, devido a perdas viscosas, produz uma queda de pressão em função do fluxo que o atravessa e possui uma função de transferência com componentes essencialmente reais na faixa de freqüência de interesse.

Qual é o análogo elétrico de um elemento passivo com fluido que, devido à compressibilidade do fluido, produz uma pressão atrasada em relação ao fluxo por essencialmente 90°?

Qual é o análogo elétrico de um elemento passivo com fluido que, devido à inércia do fluido, possui uma queda de pressão avançada em relação ao fluxo por essencialmente 90°?

6

Análise de Resposta Transitória

6.1 INTRODUÇÃO

Afirmamos no Cap. 4 que o primeiro passo na análise de um sistema de controle era derivar um modelo matemático para o sistema. Uma vez obtido tal modelo, existem vários métodos disponíveis para a análise do desempenho do sistema.

Na prática, o sinal de entrada em um sistema de controle não é conhecido *a priori* mas é de caráter aleatório, e a entrada instantânea não pode ser expressa analiticamente. Apenas em alguns casos especiais se conhece o sinal de entrada *a priori* e se pode expressá-lo analiticamente ou por curvas, como no caso do controle automático de ferramentas de corte.

Na análise e projeto de sistemas de controle, devemos ter uma base para comparar o desempenho de vários sistemas de controle. Esta base pode ser obtida especificando-se sinais de teste de entrada particulares e comparando-se as respostas de vários sistemas a estes sinais de entrada.

Muitos critérios de projeto são baseados em tais sinais ou na resposta de sistemas a mudanças nas condições iniciais (sem utilizar qualquer sinal de teste). O uso de sinais de teste pode ser justificado pela correlação que existe entre as características de um sistema para um sinal de entrada de teste típico e a capacidade do sistema para responder aos sinais de entrada reais.

Sinais de teste típicos. Os sinais de entrada de teste comumente usados são as funções-degrau, funções-rampa, funções-aceleração, funções-impulso, funções senoidais etc. Com estes sinais de teste, tanto análises matemáticas como experimentais de sistemas de controle podem ser feitas com facilidade, pois estes sinais são funções muito simples do tempo.

A determinação de qual ou quais destes sinais de entrada típicos devem ser usados para analisar características do sistema depende da forma da entrada a que o sistema será sujeito mais freqüentemente durante operação normal. Se as entradas para um sistema de controle são funções que variam gradativamente com o tempo, então uma função rampa de tempo pode ser um bom sinal de teste; e para um

sistema sujeito a entradas tipo choque, uma função impulso pode ser o melhor. Uma vez projetado um sistema de controle baseado em sinais de teste, normalmente o desempenho do sistema para entradas reais é satisfatório. O uso de tais sinais de teste nos permite comparar o desempenho de todos os sistemas na mesma base.

Este capítulo trata da resposta de sistemas a sinais aperiódicos (como, por exemplo, funções do tempo: degrau, rampa, aceleração e impulso). (Análise e projeto de sistemas baseados em sinais de teste senoidais são dados nos Caps. 9 e 10.)

Resposta transitória e resposta estacionária. A resposta temporal de um sistema de controle consiste em duas partes: a resposta transitória e a resposta estacionária. Entendemos por resposta transitória aquela que vai do estado inicial até o estado final. Por resposta estacionária, entendemos a maneira como a saída do sistema se comporta quando t tende a infinito.

Estabilidade absoluta, estabilidade relativa, e erro estacionário. Ao projetar um sistema de controle, devemos poder prever o comportamento dinâmico do sistema a partir do conhecimento dos componentes. A característica mais importante do comportamento dinâmico de um sistema de controle é a estabilidade absoluta, isto é, se o sistema é estável ou instável. Um sistema de controle está em equilíbrio se, na ausência de qualquer perturbação ou entrada, a saída permanece no mesmo estado. Um sistema de controle invariante no tempo e linear é estável se a saída volta ao seu estado de equilíbrio quando o sistema é sujeito a uma perturbação. Um sistema de controle, invariante no tempo e linear, é instável se uma oscilação da saída continua indefinidamente ou a saída diverge sem limite a partir de seu estado de equilíbrio quando o sistema é sujeito a uma perturbação. Em casos reais, a saída de um sistema físico pode aumentar até um certo valor mas ser limitada por "barreiras" mecânicas, ou o sistema pode parar de funcionar ou se tornar não linear após a saída ultrapassar uma certa amplitude de tal forma que as equações diferenciais lineares não são mais válidas. Não discutiremos a estabilidade de sistemas não lineares neste capítulo. (Tais análises sobre sistemas não lineares são vistas nos Caps. 11 e 15.)

Outros comportamentos (além de estabilidade absoluta) do sistema que devemos considerar cuidadosamente incluem a estabilidade relativa e o erro estacionário. Dado que um sistema de controle físico envolve armazenamento de energia, a saída do sistema, quando sujeito a uma entrada, não pode seguir a entrada imediatamente mas exibe uma resposta transitória antes que um regime permanente possa ser alcançado. A resposta transitória de um sistema de controle real muitas vezes mostra oscilações amortecidas antes de ir a um estado ou regime estacionário. Se a saída de um sistema em regime estacionário não concorda exatamente com a entrada, diz-se que o sistema apresenta erro estacionário. Este erro indica a precisão do sistema. Ao analisar um sistema de controle, devemos examinar o comportamento da resposta transitória, tal como o tempo necessário para alcançar um novo regime ou estado estacionário e o valor do erro enquanto o sistema segue um sinal de entrada, bem como o comportamento em regime estacionário.

6.2 FUNÇÕES DE RESPOSTA AO IMPULSO

Para um sistema linear, invariante no tempo, a função de transferência $G(s)$ é

$$G(s) = \frac{Y(s)}{X(s)}$$

onde $X(s)$ é a transformada de Laplace da entrada e $Y(s)$ é a transformada de Laplace da saída. Segue-se que a saída $Y(s)$ pode ser formulada como o produto de $G(s)$ e $X(s)$, ou

$$Y(s) = G(s)X(s) \tag{6.1}$$

Note que multiplicação no domínio complexo é equivalente à convolução no domínio do tempo, e portanto a transformada inversa de Laplace da Eq. (6.1) é dada pela seguinte integral de convolução:

$$\begin{aligned} y(t) &= \int_0^t x(\tau)g(t-\tau)\,d\tau \\ &= \int_0^t g(\tau)x(t-\tau)\,d\tau \end{aligned}$$

onde

$$g(t) = x(t) = 0 \qquad \text{para } t < 0$$

Funções de resposta ao impulso. Considere a saída (resposta) de um sistema a uma entrada de impulso unitário quando as condições iniciais são nulas. Como a transformada de Laplace da função impulso unitário é a unidade, a transformada de Laplace da saída do sistema é simplesmente

$$Y(s) = G(s) \tag{6.2}$$

A transformada inversa de Laplace da saída do sistema, dada pela Eq. (6.2), é a função resposta ao impulso, ou

$$y(t) = g(t) = \text{função resposta ao impulso}$$

Esta função também é chamada a função "peso" do sistema.

A função resposta ao impulso $g(t)$ é portanto a resposta de um sistema linear a uma entrada de impulso unitário quando as condições iniciais são nulas. A transformada de Laplace desta função nos dá a função de transferência. Portanto, a função de transferência e a função resposta ao impulso de um sistema linear invariante no tempo contém a mesma informação sobre a dinâmica do sistema. Desta forma, é possível obter uma informação completa sobre as características dinâmicas do sistema excitando-o com uma entrada impulsiva e medindo a resposta. Na prática, uma entrada em forma de um pulso, cuja duração é muito curta comparada com as constantes de tempo significativas do sistema, pode ser considerada como impulsiva.

Considere a resposta de um sistema de primeira-ordem para uma entrada em pulso de amplitude $1/t_1$ e duração t_1. Se a duração $0 < t < t_1$ desta entrada é suficientemente pequena comparada com a constante de tempo do sistema T, então a resposta é aproximadamente uma resposta ao impulso unitário. Um método para determinar se t_1 é suficientemente pequeno é aplicar uma entrada em pulso com amplitude $2/t_1$ e duração $t_1/2$. Se as respostas à entrada em pulso original e do pulso modificado são praticamente as mesmas, t_1 pode ser considerado suficientemente pequeno. A Fig. 6.1 mostra as curvas de resposta de um sistema de primeira-ordem a entradas em pulso e a entrada em impulso. Observe que, se $t_1 < 0.1T$, a resposta do sistema é praticamente idêntica à resposta ao impulso unitário.

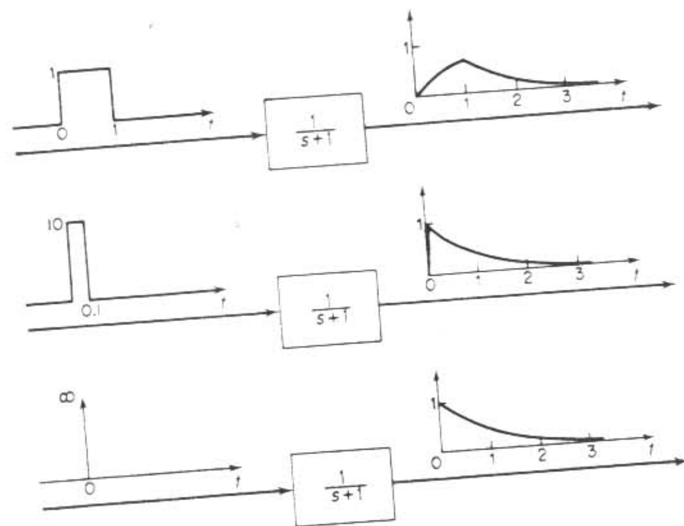


Fig. 6.1 Curvas de resposta para um sistema de primeira-ordem sujeito a entradas em pulso e impulso.

Integrais de convolução. Na prática podemos aproximar uma função impulso por uma função pulso de amplitude grande e de duração pequena, cuja área é igual à magnitude, ou área, da função impulso. Se a função de entrada $x(t)$ é uma função em pulso com amplitude $1/t_1$ e duração t_1, então a integral de convolução

$$y(t) = \int_0^t x(\tau)g(t-\tau)\,d\tau$$

se torna

$$\begin{aligned} y(t) &= \int_0^{t_1} \frac{1}{t_1} g(t-\tau)\,d\tau \\ &\doteq g(t) \qquad \text{para } t_1 \text{ suficientemente pequeno} \end{aligned}$$

Portanto, a resposta de um sistema para um pulso de grande amplitude e pequena duração tendo área unitária é praticamente igual à função resposta ao impulso.

Aproximação para integrais de convolução. Considere o sistema visto na Fig. 6.2. Suponha que a função resposta ao impulso é $g(t)$. {É claro que a função de transferência do sistema é $\mathscr{L}[g(t)] = G(s)$}. A entrada $x(t)$ se inicia em $t = 0$ e dura até $t = t_1$. Achemos a resposta deste sistema para $x(t)$ usando uma aproximação para a integral de convolução:

$$y(t) = \int_0^t x(t)g(t-\tau)\,d\tau$$

pela soma das respostas a N funções em pulso.

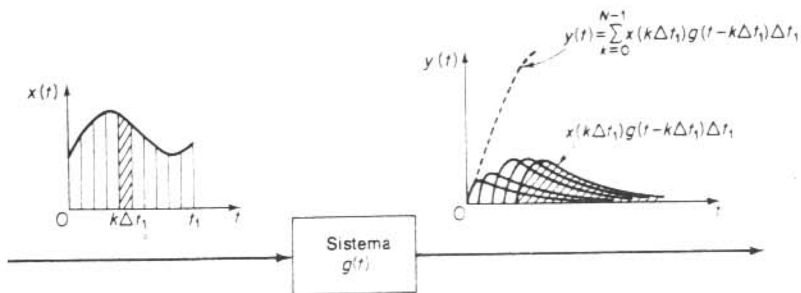


Fig. 6.2 Saída do sistema como uma somatória de convolução.

A entrada $x(t)$ pode ser aproximada por uma seqüência de N funções em pulso cuja duração é Δt_1, onde $\Delta t_1 = t_1/N$. Se Δt_1 é suficientemente pequeno quando comparado com a menor constante de tempo do sistema, então o k-ésimo pulso pode ser considerado como um impulso cuja magnitude é a área $x(k\Delta t_1)\,\Delta t_1$. A resposta ao k-ésimo pulso é então

$$x(k\,\Delta t_1)\,\Delta t_1\,g(t - k\,\Delta t_1)$$

que é o produto da área do impulso e a função resposta ao impulso atrasada por $k\Delta t_1$.

Como o sistema em consideração é linear, o princípio da superposição se aplica. Portanto a resposta $y(t)$ do sistema à seqüência de N funções em pulso é dada pela soma de convolução:

$$y(t) = \sum_{k=0}^{N-1} x(k\,\Delta t_1)g(t - k\,\Delta t_1)\,\Delta t_1 \tag{6.3}$$

onde

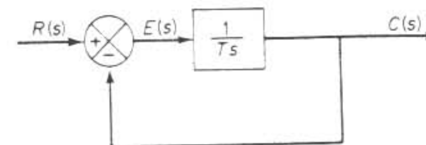
$$g(\tau) = 0 \qquad \text{para } \tau < 0$$

A Eq. (6.3) fornece a resposta no instante t. Note que como $g(\tau) = 0$ para $\tau < 0$, a resposta não precede a entrada. Portanto,

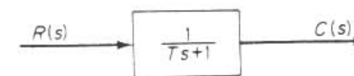
$$\begin{aligned}
y(0 \le t < \Delta t_1) &= x(0)g(t)\,\Delta t_1 \\
y(\Delta t_1 \le t < 2\,\Delta t_1) &= [x(0)g(t) - x(\Delta t_1)g(t - \Delta t_1)]\,\Delta t_1 \\
y(2\,\Delta t_1 \le t < 3\,\Delta t_1) &= [x(0)g(t) - x(\Delta t_1)g(t - \Delta t_1) \\
&\quad + x(2\,\Delta t_1)g(t - 2\,\Delta t_1)]\,\Delta t_1 \\
&\cdots \\
y(N\,\Delta t_1 - \Delta t_1 \le t \le N\,\Delta t_1) &= \left[\sum_{k=0}^{N-1} x(k\,\Delta t_1)g(t - k\,\Delta t_1)\right]\Delta t_1
\end{aligned}$$

6.3 SISTEMAS DE PRIMEIRA-ORDEM

Considere o sistema de primeira-ordem visto na Fig. 6.3(a).



(a)



(b)

Fig. 6.3 (a) Diagrama de blocos de um sistema de primeira-ordem; (b) diagrama de blocos simplificado.

Fisicamente, este sistema pode representar um circuito R-C, um sistema térmico etc. Um diagrama de blocos simplificado é visto na Fig. 6.3(b). A relação entrada-saída é dada por

$$\frac{C(s)}{R(s)} = \frac{1}{Ts + 1} \tag{6.4}$$

A seguir, analisaremos as respostas do sistema a entradas como degrau unitário, rampa unitária, e função impulso unitário. As condições iniciais serão nulas.

Note que todos os sistemas com a mesma função de transferência irão exibir a mesma saída em resposta à mesma entrada. Para qualquer sistema físico dado, pode-se dar uma interpretação física à resposta matemática.

Resposta a degrau unitário de sistemas de primeira-ordem. Como a transformada de Laplace da função degrau unitário é $1/s$, substituindo $R(s) = 1/s$ na Eq. (6.4), obtemos

$$C(s) = \frac{1}{Ts + 1}\,\frac{1}{s}$$

Expandindo $C(s)$ em frações parciais, temos

$$C(s) = \frac{1}{s} - \frac{T}{Ts + 1} \tag{6.5}$$

Tomando a transformada inversa de Laplace da Eq. (6.5), obtemos

$$c(t) = 1 - e^{-t/T} \qquad (t \ge 0) \tag{6.6}$$

A Eq. (6.6) diz que, inicialmente, a saída $c(t)$ é nula e finalmente se torna unitária. Uma das características importantes desta curva de resposta exponencial $c(t)$ é que em $t = T$ o valor de $c(t)$ é 0,632, ou a resposta $c(t)$ alcançou 63,2% de sua variação total. Isto pode ser visto facilmente substituindo-se $t = T$ em $c(t)$. Ou seja,

$$c(T) = 1 - e^{-1} = 0{,}632$$

Sabe-se que T é a constante de tempo do sistema. Quanto menor for a constante de tempo, mais rápida será a resposta do sistema. Outra característica importante da curva de resposta exponencial é que a inclinação da linha tangente em $t = 0$ é $1/T$, pois

$$\frac{dc}{dt} = \frac{1}{T} e^{-t/T} \bigg|_{t=0} = \frac{1}{T} \tag{6.7}$$

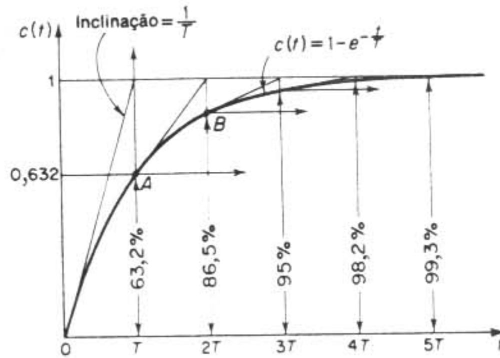


Fig. 6.4 Curva de resposta exponencial.

A saída alcançaria o valor final em $t = T$ caso se mantivesse a sua velocidade inicial de resposta. Da Eq. (6.7) vemos que a inclinação da curva de resposta $c(t)$ decresce monotonicamente de $1/T$ em $t = 0$ para zero em $t = \infty$.

A curva de resposta exponencial $c(t)$ dada pela Eq. (6.6) é vista na Fig. 6.4. Em uma constante de tempo, a curva de resposta exponencial foi de 0 a 63,2% do valor final. Em duas constantes de tempo, a resposta alcança 86,5% do valor final. Em $t = 3T$, $4T$ e $5T$, a resposta alcança 95, 98,2 e 99,3%, respectivamente, do valor final. Portanto, para $t \geq 4T$, a resposta permanece dentro de 2% do valor final. Como visto da Eq. (6.6), o regime estacionário é alcançado matematicamente somente após um tempo infinito. Na prática, entretanto, uma estimativa razoável do tempo de resposta é o tempo que a curva de resposta necessita para alcançar a linha de 2% do valor final, ou quatro constantes de tempo.

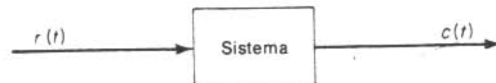


Fig. 6.5 Um sistema geral.

Considere o sistema visto na Fig. 6.5. Para determinar experimentalmente se o sistema é ou não de primeira-ordem, faça o gráfico da curva log $|c(t) - c(\infty)|$, onde $c(t)$ é a saída do sistema, em função de t. Se a curva é uma reta, o sistema é de primeira-ordem. A constante de tempo T pode ser lida do gráfico como sendo o tempo T que satisfaz a seguinte equação:

$$c(T) - c(\infty) = 0{,}368\,[c(0) - c(\infty)]$$

Note que ao invés de fazer o gráfico de log $|c(t) - c(\infty)|$ em função de t, é conveniente fazer o gráfico de $|c(t) - c(\infty)|/|c(0) - c(\infty)|$ em função de t em papel semilogarítmico, como visto na Fig. 6.6.

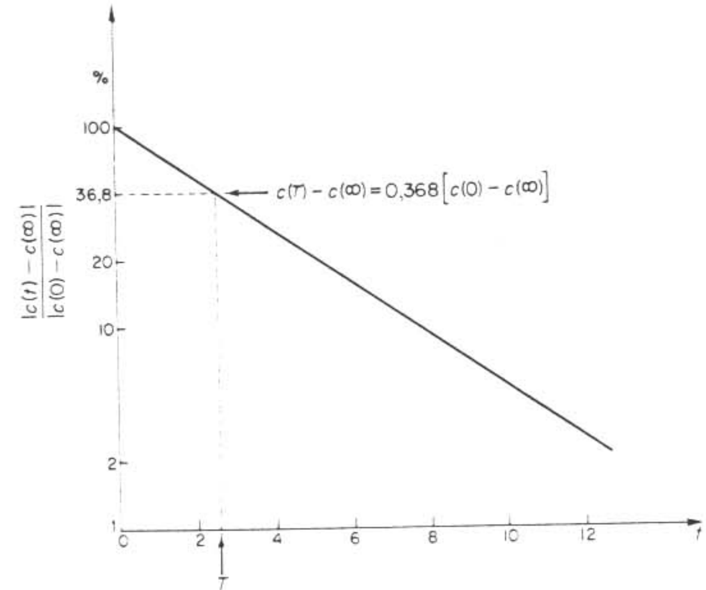


Fig. 6.6 Gráfico de $|c(t) - c(\infty)|/|c(0) - c(\infty)|$ em função de t em papel semilog.

Resposta a rampa unitária de sistemas de primeira-ordem. Como a transformada de Laplace da função rampa unitária é $1/s^2$, obtemos a saída do sistema da Fig. 6.3(a) como

$$C(s) = \frac{1}{Ts + 1} \frac{1}{s^2}$$

Expandindo $C(s)$ em frações parciais, temos

$$C(s) = \frac{1}{s^2} - \frac{T}{s} + \frac{T^2}{Ts + 1} \tag{6.8}$$

Tomando a transformada inversa de Laplace da Eq. (6.8), obtemos

$$c(t) = t - T + Te^{-t/T} \qquad (t \geq 0)$$

O sinal de erro $e(t)$ é então

$$e(t) = r(t) - c(t)$$
$$= T(1 - e^{-t/T})$$

Quando t tende a infinito, $e^{-t/T}$ tende a zero, e portanto o sinal de erro $e(t)$ tende a T ou

$$e(\infty) = T$$

A entrada rampa unitária e a saída do sistema são vistas na Fig. 6.7. O erro ao seguir a entrada rampa unitária é igual a T para t suficientemente grande. Quanto menor a constante de tempo T, menor é o erro estacionário ao seguir uma entrada em rampa.

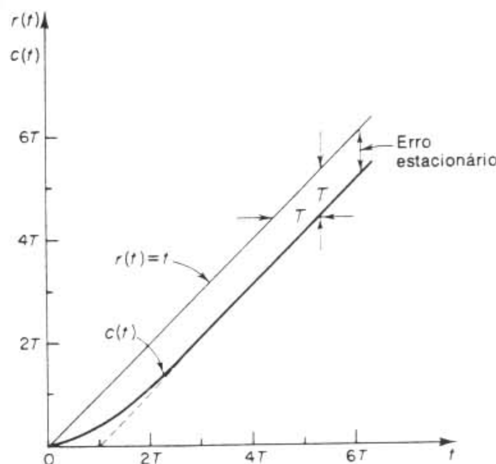


Fig. 6.7 Resposta a rampa unitária do sistema visto na Fig. 6.3(a).

Resposta ao impulso unitário de sistemas de primeira-ordem. Para a entrada impulso unitário $R(s) = 1$, e a saída do sistema da Fig. 6.3(a) pode ser obtida como

$$C(s) = \frac{1}{Ts + 1}$$

ou

$$c(t) = \frac{1}{T}e^{-t/T} \qquad (t \geq 0) \tag{6.9}$$

A curva de resposta dada pela Eq. (6.9) é vista na Fig. 6.8.

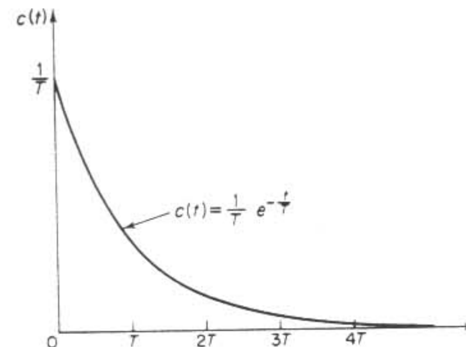


Fig. 6.8 Resposta a impulso unitário do sistema visto na Fig. 6.3(a).

Uma propriedade importante de sistemas lineares invariantes no tempo. Na análise vista acima, mostramos que para a entrada rampa unitária, a saída $c(t)$ é

$$c(t) = t - T + Te^{-t/T} \qquad (t \geq 0)$$

Para a entrada degrau unitário, que é a derivada da entrada rampa unitária, a saída $c(t)$ é

$$c(t) = 1 - e^{-t/T} \qquad (t \geq 0)$$

Finalmente, para a entrada impulso unitário, que é a derivada da entrada degrau unitário, a saída $c(t)$ é

$$c(t) = \frac{1}{T}e^{-t/T} \qquad (t \geq 0)$$

A comparação da resposta do sistema a estas três entradas mostra claramente que a resposta à derivada de um sinal de entrada pode ser obtida diferenciando a resposta do sistema para o sinal original. Também pode ser visto que a resposta à integral do sinal original pode ser obtida integrando a resposta do sistema ao sinal original e determinando as constantes de integração a partir da condição inicial de saída zero. Como mencionado no Cap. 4, esta é a propriedade de sistemas lineares invariantes no tempo. Sistemas variantes no tempo e sistemas não lineares não possuem esta propriedade.

6.4 SISTEMAS DE SEGUNDA-ORDEM

Nesta seção, vamos primeiro obter a resposta de um sistema de controle de segunda-ordem específico a uma entrada em degrau e depois estender a análise para a solução de outros sistemas de segunda-ordem.

Um servomecanismo. Considere o servomecanismo visto na Fig. 6.9(a). O objetivo deste sistema é controlar a posição da carga mecânica de acordo com a posição de referência. A operação deste sistema é como segue: Um par de potenciômetros age como um dispositivo que mede erro. Eles convertem as posições de

entrada e saída em sinais elétricos proporcionais. O sinal de entrada de comando determina a posição angular r do braço móvel do potenciômetro de entrada. A posição angular r é a entrada de referência para o sistema, e o potencial elétrico do braço é proporcional à posição angular do braço. A posição do eixo de saída determina a posição angular c do braço móvel do potenciômetro de saída. A diferença de potencial $e_r - e_c = e$ é o sinal de erro, onde e_r é proporcional a r e e_c é proporcional a c; ou seja, $e_r = K_0 r$ e $e_c = K_0 c$, onde K_0 é uma constante de proporcionalidade. O sinal de erro que aparece nos terminais do potenciômetro é amplificado pelo amplificador cuja constante de ganho é K_1. A tensão de saída deste amplificador é aplicada ao circuito de armadura do motor dc. (O amplificador deve ter uma impedância de entrada muito alta porque os potenciômetros são essencialmente circuitos de alta impedância e não toleram drenagem de corrente. Ao mesmo tempo, o amplificador deve ter impedância de saída baixa pois ele alimenta o circuito de armadura do motor.) Uma tensão fixa é aplicada ao enrolamento de campo. Se existe um erro, o motor desenvolve um torque para virar a carga de saída de tal forma a reduzir o erro para zero. Para corrente de campo constante, o torque desenvolvido pelo motor é

$$T = K_2 i_a$$

onde K_2 é a constante de torque do motor e i_a é a corrente de armadura. Para o circuito de armadura,

$$L_a \frac{di_a}{dt} + R_a i_a + K_3 \frac{d\theta}{dt} = K_1 e \tag{6.10}$$

onde K_3 é a força contra eletromotriz do motor e θ é o deslocamento angular do eixo do motor. A equação para equilíbrio de torque é

$$J_0 \frac{d^2\theta}{dt^2} + f_0 \frac{d\theta}{dt} = T = K_2 i_a \tag{6.11}$$

onde J_0 é a inércia da combinação do motor, carga, e engrenagens referidos ao eixo do motor e f_0 é o coeficiente de atrito viscoso da combinação do motor, carga e engrenagens referidos ao eixo do motor. A função de transferência entre o deslocamento angular do eixo do motor e o sinal de erro é obtida das Eqs. (6.10) e (6.11) como segue:

$$\frac{\Theta(s)}{E(s)} = \frac{K_1 K_2}{s(L_a s + R_a)(J_0 s + f_0) + K_2 K_3 s} \tag{6.12}$$

onde $\Theta(s) = \mathscr{L}[\theta(t)]$ e $E(s) = \mathscr{L}[e(t)]$. Supomos que a relação das engrenagens é tal que o eixo de saída gira n vezes para cada revolução do eixo do motor. Portanto,

$$C(s) = n\Theta(s) \tag{6.13}$$

onde $C(s) = \mathscr{L}[c(t)]$ e $c(t)$ é o deslocamento angular do eixo de saída. A relação entre $E(s)$, $R(s)$ e $C(s)$ é

$$E(s) = K_0[R(s) - C(s)] \tag{6.14}$$

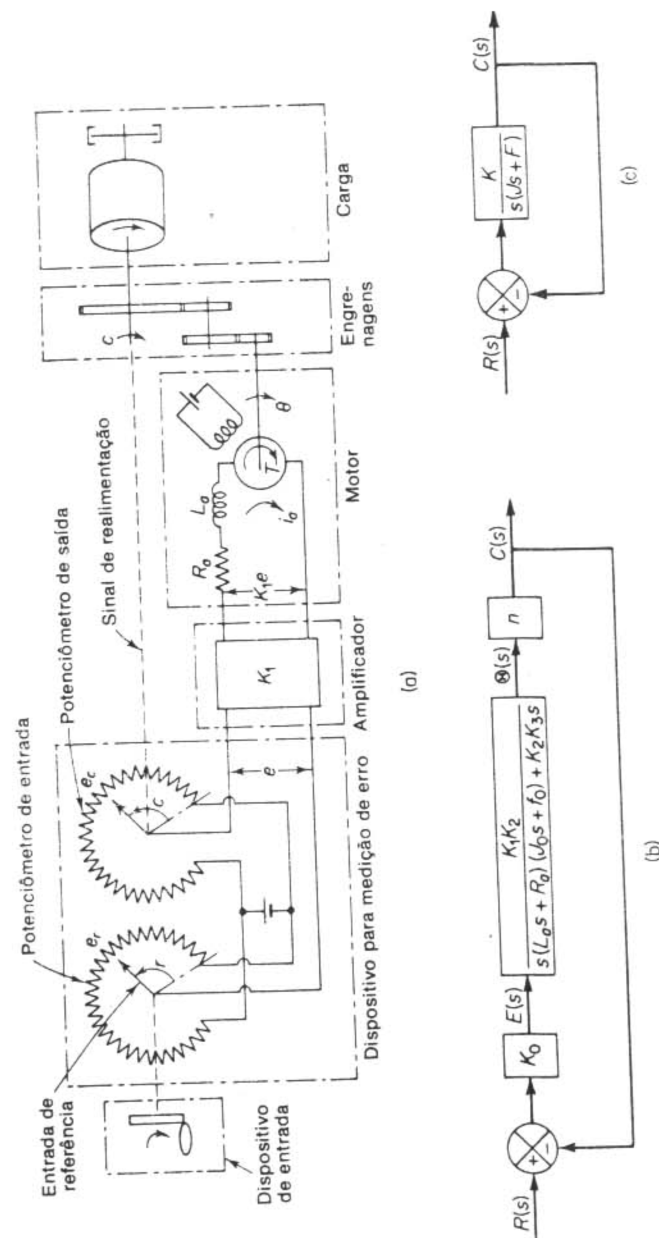


Fig. 6.9 (a) Diagrama esquemático de um servomecanismo; (b) e (c) são diagramas de blocos.

onde $R(s) = \mathcal{L}[r(t)]$. O diagrama de blocos deste sistema pode ser construído a partir das Eqs. (6.12), (6.13) e (6.14), como visto na Fig. 6.9(b). A função de transferência do elo direto deste sistema é

$$G(s) = \frac{K_0K_1K_2n}{s[(L_as + R_a)(J_0s + f_0) + K_2K_3]}$$

Como L_a é normalmente pequeno, ele pode ser desprezado, e a função de transferência de elo direto se torna

$$G(s) = \frac{K_0K_1K_2n}{s[R_a(J_0s + f_0) + K_2K_3]} = \frac{K_0K_1K_2n/R_a}{J_0s^2 + \left(f_0 + \dfrac{K_2K_3}{R_a}\right)s} \tag{6.15}$$

O termo $[f_0 + (K_2K_3/R_a)]s$ indica que a força contra eletromotriz do motor tem o efeito de aumentar o atrito viscoso do sistema. A inércia J_0 e o coeficiente de atrito viscoso $f_0 + (K_2K_3/R_a)$ são referidos ao eixo do motor. Quando J_0 e $f_0 + (K_2K_3/R_a)$ são multiplicados por $1/n^2$, a inércia e o coeficiente de atrito viscoso são expressos em termos do eixo de saída. Introduzindo novos parâmetros definidos por

$J = J_0/n^2$ = momento de inércia referido ao eixo de saída
$F = [f_0 + (K_2K_3/R_a)]/n^2$ = coeficiente de atrito viscoso referido ao eixo de saída
$K = K_0K_1K_2/nR_a$

podemos simplificar a função de transferência $G(s)$ dada pela Eq. (6.15), resultando

$$G(s) = \frac{K}{Js^2 + Fs}$$

O diagrama de blocos do sistema visto na Fig. 6.9(b) é então simplificado como visto na Fig. 6.9(c).

A seguir, investigaremos as respostas dinâmicas deste sistema para entradas degrau unitário, rampa unitária e impulso unitário.

Resposta a degrau de sistemas de segunda-ordem. A função de transferência de malha-fechada do sistema visto na Fig. 6.9(c) é

$$\frac{C(s)}{R(s)} = \frac{K}{Js^2 + Fs + K} = \frac{\dfrac{K}{J}}{\left[s + \dfrac{F}{2J} + \sqrt{\left(\dfrac{F}{2J}\right)^2 - \dfrac{K}{J}}\right]\left[s + \dfrac{F}{2J} - \sqrt{\left(\dfrac{F}{2J}\right)^2 - \dfrac{K}{J}}\right]} \tag{6.16}$$

Os pólos de malha-fechada são complexos se $F^2 - 4JK < 0$, e são reais se $F^2 - 4JK \geq 0$. Na análise de resposta transitória, é conveniente escrever

$$\frac{K}{J} = \omega_n^2, \qquad \frac{F}{J} = 2\zeta\omega_n = 2\sigma$$

onde σ é chamada de atenuação; ω_n é a freqüência natural não amortecida; e ζ é o coeficiente de amortecimento do sistema. O coeficiente de amortecimento ζ é a relação entre o amortecimento real F e o amortecimento crítico $F_c = 2\sqrt{JK}$ ou

$$\zeta = \frac{F}{F_c} = \frac{F}{2\sqrt{JK}}$$

Com esta notação, o sistema visto na Fig. 6.9(c) pode ser modificado para aquele visto na Fig. 6.10, e a função de transferência de malha-fechada $C(s)/R(s)$ dada pela Eq. (6.16) pode ser escrita

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{6.17}$$

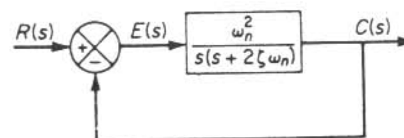


Fig. 6.10 Sistema de segunda-ordem.

O comportamento dinâmico de sistemas de segunda-ordem pode então ser descrito em termos de dois parâmetros ζ e ω_n. Se $0 < \zeta < 1$, os pólos de malha-fechada são complexos conjugados e se situam no semiplano esquerdo do plano s. O sistema então é dito subamortecido, e a resposta transitória é oscilatória. Se $\zeta = 1$, o sistema é dito criticamente amortecido. Sistemas sobreamortecidos correspondem a $\zeta > 1$. A resposta transitória de sistemas amortecidos criticamente e sobreamortecidos não oscila. Se $\zeta = 0$, a resposta transitória não decai.

Determinaremos agora a resposta do sistema visto na Fig. 6.10 para entrada degrau unitário. Consideramos três casos diferentes: o subamortecido $(0 < \zeta < 1)$, o criticamente amortecido $(\zeta = 1)$, e o sobreamortecido $(\zeta > 1)$.

(1) Caso subamortecido $(0 < \zeta < 1)$: Neste caso, $C(s)/R(s)$ pode ser escrito

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{(s + \zeta\omega_n + j\omega_d)(s + \zeta\omega_n - j\omega_d)}$$

onde $\omega_d = \omega_n\sqrt{1 - \zeta^2}$. A freqüência ω_d é chamada de freqüência natural amortecida. Para uma entrada degrau unitário, $C(s)$ pode ser escrita

$$C(s) = \frac{\omega_n^2}{(s^2 + 2\zeta\omega_n s + \omega_n^2)s} \tag{6.18}$$

A transformada inversa de Laplace da Eq. (6.18) pode ser obtida facilmente se $C(s)$ é escrita da seguinte forma:

$$C(s) = \frac{1}{s} - \frac{s + 2\zeta\omega_n}{s^2 + 2\zeta\omega_n s + \omega_n^2} = \frac{1}{s} - \frac{s + \zeta\omega_n}{(s + \zeta\omega_n)^2 + \omega_d^2} - \frac{\zeta\omega_n}{(s + \zeta\omega_n)^2 + \omega_d^2}$$

No Cap. 2 mostramos que

$$\mathscr{L}^{-1}\left[\frac{s+\zeta\omega_n}{(s+\zeta\omega_n)^2+\omega_d^2}\right] = e^{-\zeta\omega_n t}\cos\omega_d t$$

$$\mathscr{L}^{-1}\left[\frac{\omega_d}{(s+\zeta\omega_n)^2+\omega_d^2}\right] = e^{-\zeta\omega_n t}\,\mathrm{sen}\,\omega_d t$$

Portanto a transformada inversa de Laplace da Eq. (6.18) é obtida como segue:

$$\begin{aligned}\mathscr{L}^{-1}[C(s)] &= c(t)\\ &= 1 - e^{-\zeta\omega_n t}\left(\cos\omega_d t + \frac{\zeta}{\sqrt{1-\zeta^2}}\,\mathrm{sen}\,\omega_d t\right)\\ &= 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\,\mathrm{sen}\left(\omega_d t + \tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}\right) \qquad (t \geq 0)\end{aligned} \tag{6.19}$$

Este resultado pode, obviamente, ser obtido diretamente usando-se uma tabela de transformadas de Laplace. Da Eq. (6.19), pode-se ver que a freqüência da oscilação transitória é a freqüência natural amortecida ω_d e portanto varia com o coeficiente de amortecimento ζ. O sinal de erro para este sistema é a diferença entre a entrada e a saída, e portanto é

$$\begin{aligned}e(t) &= r(t) - c(t)\\ &= e^{-\zeta\omega_n t}\left(\cos\omega_d t + \frac{\zeta}{\sqrt{1-\zeta^2}}\,\mathrm{sen}\,\omega_d t\right) \qquad (t \geq 0)\end{aligned}$$

Este sinal de erro apresenta uma oscilação senoidal amortecida. Em regime permanente, ou em $t = \infty$, não existe erro entre a entrada e saída.

Se o coeficiente de amortecimento ζ é igual a zero, a resposta se torna não amortecida e as oscilações continuam indefinidamente. A resposta $c(t)$ para o caso de amortecimento nulo pode ser obtida substituindo-se $\zeta = 0$ na Eq. (6.19), resultando

$$c(t) = 1 - \cos\omega_n t \qquad (t \geq 0) \tag{6.20}$$

Portanto, da Eq. (6.20), vemos que ω_n representa a freqüência natural não amortecida do sistema. Isto é, ω_n é a freqüência em que o sistema oscilaria se o amortecimento fosse diminuído para zero. Se o sistema linear tem amortecimento, mesmo que só um pouco, a freqüência natural não amortecida não pode ser observada experimentalmente. A freqüência que pode ser observada é a freqüência natural amortecida ω_d, que é igual a $\omega_n\sqrt{1-\zeta^2}$. Esta freqüência é sempre menor que a freqüência natural não amortecida. Um aumento em ζ irá reduzir a freqüência natural amortecida ω_d. Se ζ é aumentado além da unidade, a resposta se torna sobreamortecida e não irá oscilar.

(2) Caso de amortecimento crítico ($\zeta = 1$): Se os dois pólos de $C(s)/R(s)$ são aproximadamente iguais, o sistema pode ser aproximado por um com amortecimento crítico.

Para uma entrada degrau unitário, $R(s) = 1/s$ e $C(s)$ pode ser escrita

$$C(s) = \frac{\omega_n^2}{(s+\omega_n)^2 s} \tag{6.21}$$

A transformada inversa de Laplace da Eq. (6.21) pode ser determinada como sendo

$$c(t) = 1 - e^{-\omega_n t}(1+\omega_n t) \qquad (t \geq 0) \tag{6.22}$$

Este resultado pode obviamente ser obtido fazendo-se ζ tender a 1 (um) na Eq. (6.19) e usando-se o seguinte limite:

$$\lim_{\zeta\to 1}\frac{\mathrm{sen}\,\omega_d t}{\sqrt{1-\zeta^2}} = \lim_{\zeta\to 1}\frac{\mathrm{sen}\,\omega_n\sqrt{1-\zeta^2}\,t}{\sqrt{1-\zeta^2}} = \omega_n t$$

(3) Caso sobreamortecido ($\zeta > 1$): Neste caso, os dois pólos de $C(s)/R(s)$ são reais, negativos e distintos. Para uma entrada degrau unitário, $R(s) = 1/s$ e $C(s)$ pode ser escrita

$$C(s) = \frac{\omega_n^2}{(s+\zeta\omega_n+\omega_n\sqrt{\zeta^2-1})(s+\zeta\omega_n-\omega_n\sqrt{\zeta^2-1})s} \tag{6.23}$$

A transformada inversa de Laplace da Eq. (6.23) é

$$\begin{aligned}c(t) &= 1 + \frac{1}{2\sqrt{\zeta^2-1}(\zeta+\sqrt{\zeta^2-1})}e^{-(\zeta+\sqrt{\zeta^2-1})\omega_n t}\\ &\quad - \frac{1}{2\sqrt{\zeta^2-1}(\zeta-\sqrt{\zeta^2-1})}e^{-(\zeta-\sqrt{\zeta^2-1})\omega_n t}\\ &= 1 + \frac{\omega_n}{2\sqrt{\zeta^2-1}}\left(\frac{e^{-s_1 t}}{s_1} - \frac{e^{-s_2 t}}{s_2}\right) \qquad (t \geq 0)\end{aligned} \tag{6.24}$$

onde $s_1 = (\zeta + \sqrt{\zeta^2-1})\omega_n$ e $s_2 = (\zeta - \sqrt{\zeta^2-1})\omega_n$. Portanto, a resposta $c(t)$ inclui dois termos de exponencial decrescente.

Quando ζ é consideravelmente maior que a unidade, uma das duas exponenciais decrescentes decai mais rapidamente que a outra, de tal forma que o termo da exponencial mais rápida (que corresponde a uma constante de tempo menor) pode ser desprezado. Isto é, se $-s_2$ está localizado muito mais perto do eixo $j\omega$ do que $-s_1$ (o que significa $|s_2| << |s_1|$), então, para se obter uma solução aproximada podemos desprezar $-s_1$. Isto é válido porque o efeito de $-s_1$ na resposta é muito menor que o de $-s_2$, pois o termo contendo s_1 na Eq. (6.24) decai muito mais rapidamente do que o termo contendo s_2. Uma vez que o termo de exponencial mais rápida desaparece, a resposta é similar à de um sistema de primeira-ordem, e $C(s)/R(s)$ pode ser aproximado por

$$\frac{C(s)}{R(s)} = \frac{\zeta\omega_n - \omega_n\sqrt{\zeta^2-1}}{s+\zeta\omega_n-\omega_n\sqrt{\zeta^2-1}} = \frac{s_2}{s+s_2}$$

Esta forma aproximada é uma conseqüência direta do fato de que os valores inicial e final tanto do $C(s)/R(s)$ original como da aproximação são iguais.

Com a função de transferência aproximada $C(s)/R(s)$, a resposta a degrau unitário pode ser obtida como

$$C(s) = \frac{\zeta\omega_n - \omega_n\sqrt{\zeta^2-1}}{(s+\zeta\omega_n-\omega_n\sqrt{\zeta^2-1})s}$$

A resposta temporal $c(t)$ é então

$$c(t) = 1 - e^{-(\zeta - \sqrt{\zeta^2 - 1})\omega_n t} \qquad (t \geq 0)$$

Isto fornece uma resposta aproximada ao degrau unitário quando um dos pólos de $C(s)/R(s)$ pode ser desprezado. Um exemplo da resposta temporal aproximada $c(t)$ com $\zeta = 2, \omega_n = 1$ é visto na Fig. 6.11, junto com a solução exata para $c(t)$. A solução aproximada é

$$c(t) = 1 - e^{-0,27t} \qquad (t \geq 0)$$

e a solução exata para este caso é

$$c(t) = 1 + 0,077e^{-3,73t} - 1,077e^{-0,27t} \qquad (t \geq 0)$$

A principal diferença entre as curvas de resposta exata e aproximada reside na parte inicial das curvas de resposta.

Uma família de curvas $c(t)$ com vários valores de ζ é vista na Fig. 6.12, onde a abscissa é a variável adimensional $\omega_n t$. As curvas são funções somente de ζ. Estas curvas são obtidas das Eqs. (6.19), (6.22) e (6.24). O sistema descrito por estas equações está inicialmente em repouso.

Note que dois sistemas de segunda-ordem com o mesmo ζ mas diferentes ω_n terão o mesmo sobre-sinal e o mesmo padrão oscilatório. Diz-se que tais sistemas possuem a mesma estabilidade relativa.

É importante notar que para sistemas de segunda-ordem cujas funções de transferência de malha-fechada são diferentes daquela dada pela Eq. (6.17), as

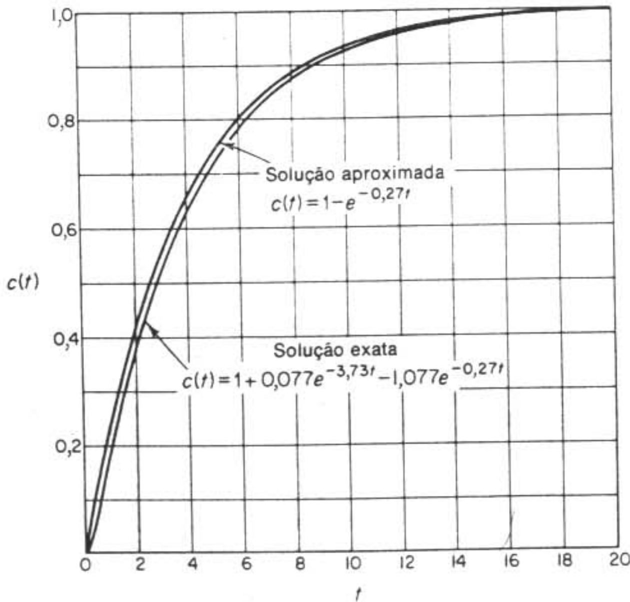


Fig. 6.11 Curvas de resposta a degrau unitário para o sistema visto na Fig. 6.10 (caso sobreamortecido.).

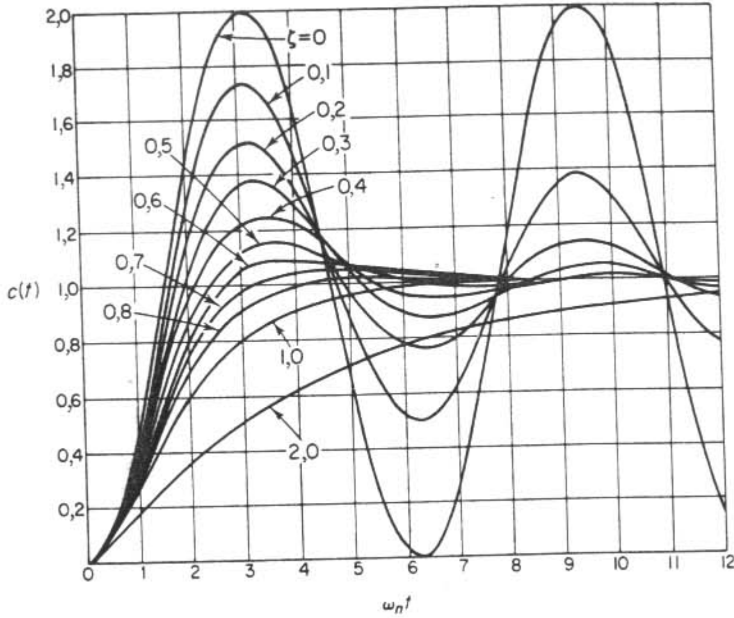


Fig. 6.12 Curvas de resposta a degrau unitário para o sistema visto na Fig. 6.10.

curvas de resposta ao degrau podem ter um aspecto bem diferente daquelas mostradas na Fig. 6.12.

Da Fig. 6.12, vemos que um sistema subamortecido com ζ entre 0,5 e 0,8 chega perto do valor final mais rapidamente que um sistema com amortecimento crítico ou um sistema sobreamortecido. Entre os sistemas que respondem sem oscilação, um sistema com amortecimento crítico apresenta a resposta mais rápida. Um sistema sobreamortecido sempre é lento ao responder a quaisquer entradas.

Definições de especificações de resposta transitória. Em muitos casos práticos, as características de desempenho desejadas de sistemas de controle são especificadas em termos e grandezas no domínio do tempo. Sistemas com armazenamento de energia não podem responder instantaneamente e terão respostas transitórias sempre que sujeitos a entradas ou perturbações.

Freqüentemente, as características de desempenho de um sistema de controle são especificadas em termos da resposta transitória para uma entrada degrau unitário, pois esta entrada é fácil de se gerar e é suficientemente severa. (Se a resposta a uma entrada em degrau é conhecida, é matematicamente possível computar a resposta para qualquer entrada.)

A resposta transitória de um sistema para uma entrada degrau unitário depende das condições iniciais. Para conveniência na comparação de respostas transitórias de vários sistemas, costuma-se usar a condição inicial padrão de que o sistema está inicialmente em repouso com a saída e todas suas derivadas nulas. Desta forma as características de resposta podem ser facilmente comparadas.

A resposta transitória de um sistema de controle na prática geralmente apresenta oscilações amortecidas antes de alcançar o estado ou regime estacionário. Ao

especificar as características de resposta transitória de um sistema de controle para uma entrada degrau unitário, é comum se especificar o seguinte:*

1. tempo de atraso, t_d
2. tempo de subida, t_r
3. instante do pico, t_p
4. sobre-sinal máximo, M_p
5. tempo de acomodação, t_s

Estas especificações são definidas a seguir e mostradas graficamente na Fig. 6.13.

1. Tempo de atraso, t_d: o tempo de atraso é o tempo necessário para a resposta alcançar pela primeira vez a metade do valor final.
2. Tempo de subida, t_r: o tempo de subida é o tempo necessário para a resposta passar de 10% a 90%, 5% a 95%, ou 0% a 100% do seu valor final. Para sistemas de segunda-ordem subamortecidos, usa-se normalmente o tempo de subida de 0 a 100%. Para sistemas sobreamortecidos, normalmente se usa o tempo de subida de 10 a 90%.
3. Instante do pico, t_p: o instante do pico é o tempo necessário para a resposta alcançar o primeiro pico do sobre-sinal.
4. Sobre-sinal máximo (percentual), M_p: o sobre-sinal máximo é o máximo valor de pico da curva de resposta medido a partir do valor unitário (um). Se o valor final de regime estacionário da resposta difere da unidade, então comumente se usa o máximo sobre-sinal percentual. É definido por

$$\text{Sobre-sinal máximo percentual} = \frac{c(t_p) - c(\infty)}{c(\infty)} \times 100\%$$

O valor do sobre-sinal máximo (percentual) indica diretamente a estabilidade relativa do sistema.

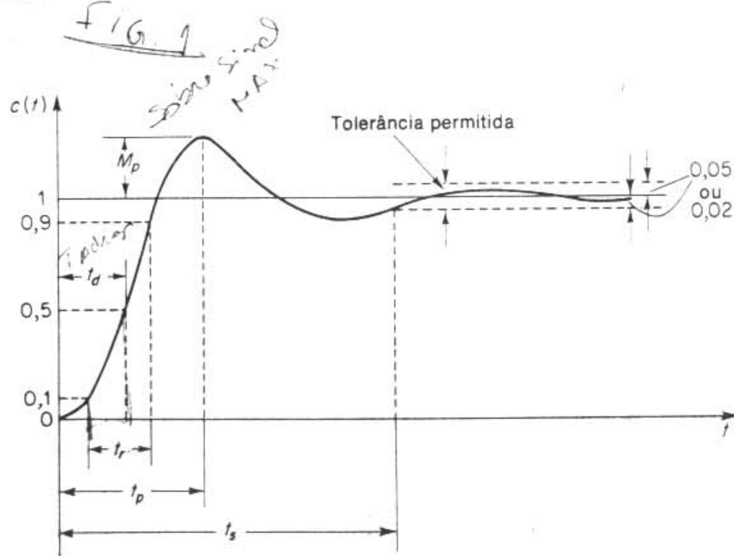


Fig. 6.13 Curva de resposta a degrau unitário mostrando t_d, t_r, t_p, M_p, e t_s.

*N. do T.: Os índices em t_d, t_r, t_p, t_s vêm do inglês *delay time*, *rise time*, *peak time*, *settling time*.

5. Tempo de acomodação, t_s: o tempo de acomodação é o tempo necessário para a curva de resposta alcançar e permanecer dentro de uma faixa em torno do valor final, faixa esta de magnitude especificada por uma porcentagem absoluta do valor final (normalmente 5% ou 2%). O tempo de acomodação está relacionado com a maior constante de tempo do sistema de controle. A escolha de que porcentagem usar no critério de erro pode ser determinada a partir dos objetivos do projeto de sistema em questão.

As especificações de domínio de tempo que acabamos de ver são bastante importantes visto que a maioria dos sistemas de controle são sistemas de domínio do tempo; isto é, eles devem apresentar respostas temporais aceitáveis. (Isto significa que o sistema de controle deve ser modificado até que a resposta transitória seja satisfatória.) Note que se especificamos os valores de t_d, t_r, t_p, t_s e M_p, então a forma da curva de resposta está virtualmente determinada. Isto pode ser visto claramente na Fig. 6.14.

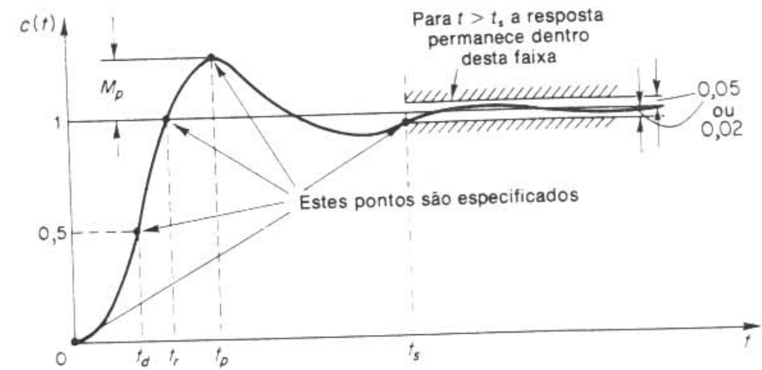


Fig. 6.14 Especificações de resposta transitória.

Note que nem todas estas especificações são aplicáveis para qualquer caso em estudo. Por exemplo, para um sistema sobreamortecido, os termos instante do pico e sobre-sinal máximo não se aplicam. (Para sistemas que apresentam erros de regime estacionário para entradas em degrau, este erro deve ser mantido dentro de um nível percentual especificado. Discussões detalhadas de erros de regime estacionário são vistas no Cap. 7.)

Alguns comentários sobre especificações de respostas transitórias. Exceto em certas aplicações onde não se podem tolerar oscilações, é desejável que a resposta transitória seja suficientemente rápida e suficientemente amortecida. Portanto, para uma desejada resposta transitória de um sistema de segunda-ordem, o coeficiente de amortecimento deve estar entre 0,4 e 0,8. Valores pequenos de ζ ($\zeta < 0{,}4$) resultam em sobre-sinal excessivo na resposta transitória, e um sistema com um valor grande de ζ($\zeta > 0{,}8$) responderá de forma lenta.

Veremos mais tarde que o sobre-sinal máximo e o tempo de subida são especificações conflitantes. Em outras palavras, não se podem minimizar sobre-sinal máximo e o tempo de subida simultaneamente. Se diminuirmos um deles, o outro necessariamente aumentará.

Sistemas de segunda-ordem e especificações de resposta transitória. A seguir, obteremos expressões para o tempo de subida, instante do pico, sobre-sinal máximo, e tempo de acomodação de sistemas de segunda-ordem dados pela Eq. (6.17). Estas expressões serão obtidas em termos de ζ e ω_n. Supõe-se que o sistema é subamortecido.

Tempo de subida t_r: Referindo-se à Eq. (6.19), obtemos o tempo de subida t_r, fazendo $c(t_r) = 1$ ou

$$c(t_r) = 1 = 1 - e^{-\zeta\omega_n t_r}\left(\cos\omega_d t_r + \frac{\zeta}{\sqrt{1-\zeta^2}}\operatorname{sen}\omega_d t_r\right) \tag{6.25}$$

Como $e^{-\zeta\omega_n t_r} \neq 0$, obtemos da Eq. (6.25) a seguinte equação:

$$\cos\omega_d t_r + \frac{\zeta}{\sqrt{1-\zeta^2}}\operatorname{sen}\omega_d t_r = 0$$

ou

$$\tan\omega_d t_r = -\frac{\sqrt{1-\zeta^2}}{\zeta} = -\frac{\omega_d}{\sigma}$$

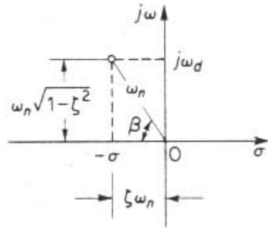


Fig. 6.15 Definição do ângulo β.

Portanto, o tempo de subida t_r é

$$t_r = \frac{1}{\omega_d}\tan^{-1}\left(\frac{\omega_d}{-\sigma}\right) = \frac{\pi - \beta}{\omega_d} \tag{6.26}$$

onde β é definido na Fig. 6.15. Pode-se ver que os valores de $\tan^{-1}(-\sqrt{1-\zeta^2}/\zeta)$ fica entre $\pi/2$ e π. Se $\zeta = 0_+$, então $\tan^{-1}(-\sqrt{1-\zeta^2}/\zeta) = \pi_+/2$; e se $\zeta = 1_-$, então $\tan^{-1}(-\sqrt{1-\zeta^2}/\zeta) = \pi_-$. Claramente, para um valor pequeno de t_r, ω_n deve ser grande.

Instante do pico t_p: Referindo-se à Eq. (6.19), podemos obter o instante do pico diferenciando $c(t)$ com respeito ao tempo e fazendo a derivada igual a zero, ou

$$\left.\frac{dc}{dt}\right|_{t=t_p} = (\operatorname{sen}\omega_d t_p)\frac{\omega_n}{\sqrt{1-\zeta^2}}e^{-\zeta\omega_n t_p} = 0$$

Isto resulta na seguinte equação:

$$\operatorname{sen}\omega_d t_p = 0$$

ou

$$\omega_d t_p = 0, \pi, 2\pi, 3\pi, \ldots$$

Como o instante do pico t_p corresponde ao primeiro pico do sobre-sinal, $\omega_d t_p = \pi$. Portanto,

$$t_p = \frac{\pi}{\omega_d} \tag{6.27}$$

O instante do pico t_p corresponde a meio ciclo da freqüência de oscilação amortecida.

Sobre-sinal máximo M_p: O sobre-sinal máximo ocorre no instante do instante do pico ou em $t = t_p = \pi/\omega_d$. Portanto, da Eq. (6.19), M_p é obtido como

$$\begin{aligned} M_p &= c(t_p) - 1 \\ &= -e^{-\zeta\omega_n(\pi/\omega_d)}\left(\cos\pi + \frac{\zeta}{\sqrt{1-\zeta^2}}\operatorname{sen}\pi\right) \\ MP &= e^{-(\sigma/\omega_d)\pi} = e^{-(\zeta/\sqrt{1-\zeta^2})\pi} \end{aligned} \tag{6.28}$$

O sobre-sinal máximo percentual é $e^{-(\sigma/\omega_d)\pi} \times 100\%$.

Tempo de acomodação t_s: Para um sistema subamortecido de segunda-ordem, a resposta transitória é obtida da Eq. (6.19) como

$$c(t) = 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\operatorname{sen}\left(\omega_d t + \tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}\right) \qquad (t \geq 0)$$

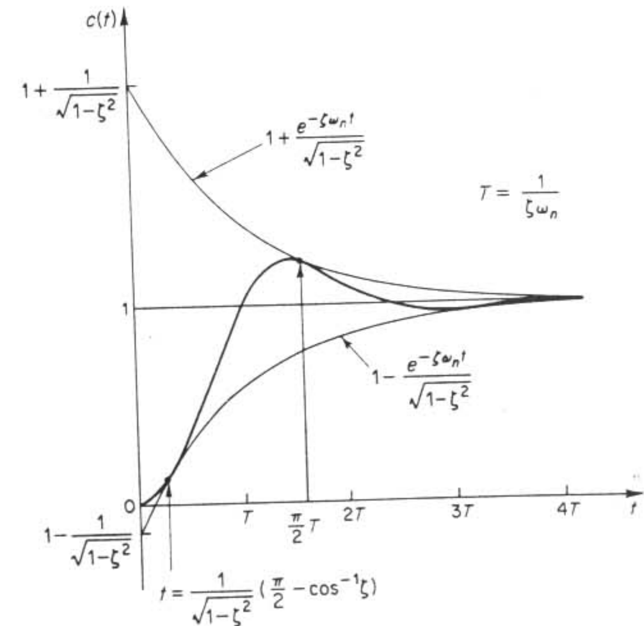


Fig. 6.16 Um par de curvas envoltórias da curva de resposta a degrau unitário do sistema visto na Fig. 6.10.

As curvas $1 \pm (e^{-\zeta\omega_n t}/\sqrt{1-\zeta^2})$ são as curvas envoltórias da resposta transitória para uma entrada degrau unitário. A curva de resposta $c(t)$ sempre permanece internamente a um par de curvas envoltórias, como visto na Fig. 6.16. A constante de tempo destas curvas envoltórias é $1/\zeta\omega_n$.

A velocidade de declínio da resposta transitória depende do valor da constante de tempo $1/\zeta\omega_n$. Para um dado ω_n, o tempo de acomodação t_s é uma função do coeficiente de amortecimento ζ. Da Fig. 6.12, vemos que para o mesmo ω_n e para uma gama de ζ entre 0 e 1, o tempo de acomodação t_s para um sistema de amortecimento muito baixo é maior do que para um sistema adequadamente amortecido. Para um sistema sobreamortecido o tempo de acomodação t_s se torna grande por causa do início lento da resposta.

O tempo de acomodação correspondendo a uma faixa de tolerância de $\pm 2\%$ ou $\pm 5\%$ pode ser medido em termos da constante de tempo $T = 1/\zeta\omega_n$, a partir das curvas da Fig. 6.12 para diferentes valores de ζ. Os resultados são mostrados na Fig. 6.17. Para $0 < \zeta < 0{,}9$, se é utilizado o critério de 2%, então t_s é aproximadamente quatro vezes a constante de tempo do sistema ou

$$t_s = 4T = \frac{4}{\sigma} = \frac{4}{\zeta\omega_n} \tag{6.29}$$

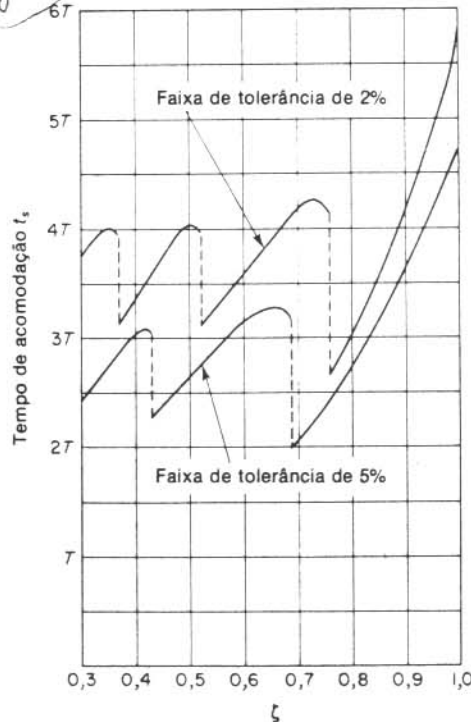


Fig. 6.17 Curvas de tempo de acomodação t_s em função de ζ.

Se é utilizado o critério de 5%, então t_s é aproximadamente três vezes a constante de tempo ou

$$t_s = 3T = \frac{3}{\sigma} = \frac{3}{\zeta\omega_n} \tag{6.30}$$

O tempo de acomodação dado pela Eq. (6.29) [ou Eq. (6.30)] tem um valor mínimo em torno de $\zeta = 0{,}76$ (ou $\zeta = 0{,}68$) e em seguida aumenta quase linearmente para valores grandes de ζ. As descontinuidades nas curvas da Fig. 6.17 acontecem porque uma variação infinitesimal no valor de ζ pode causar uma variação finita no tempo de acomodação, como visto na Fig. 6.18.

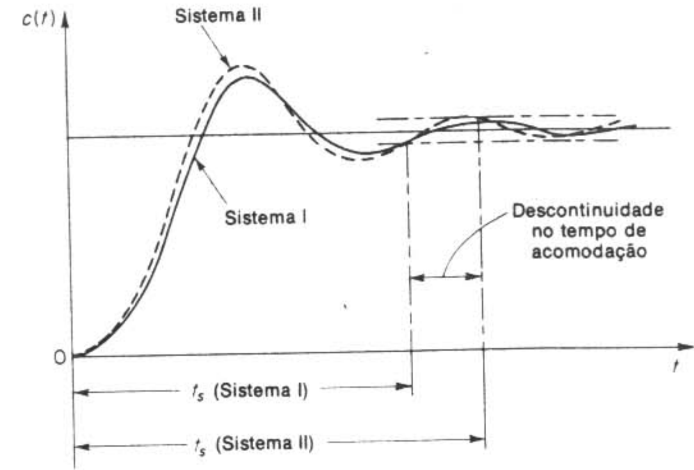


Fig. 6.18 Curvas de resposta a degrau unitário mostrando uma descontinuidade no tempo de acomodação.

Note que o tempo de acomodação é inversamente proporcional ao produto do coeficiente de amortecimento e a freqüência natural não amortecida do sistema. Como o valor de ζ é normalmente determinado a partir da especificação requerida de sobre-sinal máximo, o tempo de acomodação é determinado principalmente pela freqüência natural não amortecida ω_n. Isto significa que a duração do período transitório pode ser modificada, sem modificar o sobre-sinal máximo, através do ajuste da freqüência natural não amortecida.

Da análise anterior, é evidente que para uma resposta rápida, ω_n deve ser grande. Para limitar o sobre-sinal máximo M_p e fazer o tempo de acomodação pequeno, o coeficiente de amortecimento ζ não deve ser muito pequeno. A relação entre o sobre-sinal máximo percentual M_p e o coeficiente de amortecimento ζ é mostrada na Fig. 6.19. Note que se o coeficiente de amortecimento está entre 0,4 e 0,8, então o sobre-sinal máximo percentual para resposta a degrau está entre 25% e 2,5%.

Exemplo 6.1 Considere o sistema visto na Fig. 6.10, onde $\zeta = 0{,}6$ e $\omega_n = 5$ rad/s. Vamos obter o tempo de subida t_r, instante do pico t_p, sobre-sinal máximo M_p, e o tempo de acomodação t_s quando o sistema é sujeito a uma entrada degrau unitário.

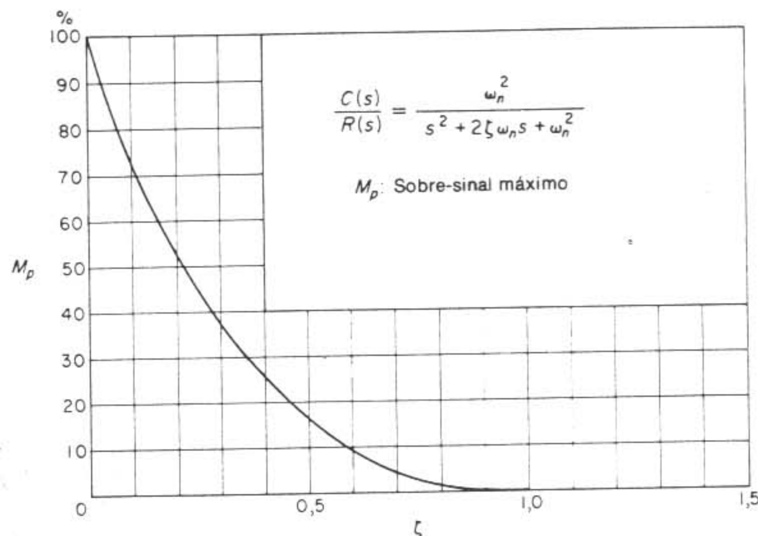


Fig. 6.19 Curva de M_p em função de ζ.

Para os valores dados de ζ e ω_n, obtemos $\omega_d = \omega_n\sqrt{1-\zeta^2} = 4$ e $\sigma = \zeta\omega_n = 3$.

Tempo de subida t_r: O tempo de subida é

$$t_r = \frac{\pi - \beta}{\omega_d} = \frac{3{,}14 - \beta}{4}$$

onde β é dado por

$$\beta = \tan^{-1}\frac{\omega_d}{\sigma} = \tan^{-1}\frac{4}{3} = 0{,}93 \text{ rad}$$

O tempo de subida t_r é então

$$t_r = \frac{3{,}14 - 0{,}93}{4} = 0{,}55 \text{ s}$$

Instante do pico t_p: O instante do pico é

$$t_p = \frac{\pi}{\omega_d} = \frac{3{,}14}{4} = 0{,}785 \text{ s}$$

Sobre-sinal máximo M_p: O sobre-sinal máximo é

$$M_p = e^{-(\sigma/\omega_d)\pi} = e^{-(3/4)\times 3{,}14} = 0{,}095$$

O sobre-sinal máximo percentual é então 9,5%.

Tempo de acomodação t_s: Para o critério de 2%, o tempo de acomodação é

$$t_s = \frac{4}{\sigma} = \frac{4}{3} = 1{,}33 \text{ s}$$

Para o critério de 5%,

$$t_s = \frac{3}{\sigma} = \frac{3}{3} = 1 \text{ s}$$

Resposta impulsiva de sistemas de segunda-ordem. Para uma entrada impulso unitário $r(t)$, a correspondente transformada de Laplace é unitária, ou $R(s) = 1$. A resposta ao impulso unitário $C(s)$ do sistema de segunda-ordem visto na Fig. 6.10 é

$$C(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

A transformada inversa de Laplace desta equação resulta na solução temporal para a resposta $c(t)$ como segue:

Para $0 \leq \zeta < 1$,

$$c(t) = \frac{\omega_n}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \operatorname{sen} \omega_n\sqrt{1-\zeta^2}\,t \qquad (t \geq 0) \tag{6.31}$$

Para $\zeta = 1$,

$$c(t) = \omega_n^2 t e^{-\omega_n t} \qquad (t \geq 0) \tag{6.32}$$

Para $\zeta > 1$,

$$c(t) = \frac{\omega_n}{2\sqrt{\zeta^2-1}} e^{-(\zeta-\sqrt{\zeta^2-1})\omega_n t} - \frac{\omega_n}{2\sqrt{\zeta^2-1}} e^{-(\zeta+\sqrt{\zeta^2-1})\omega_n t} \qquad (t \geq 0) \tag{6.33}$$

Note que podemos, sem tomar a transformada inversa de Laplace de $C(s)$, obter a resposta temporal $c(t)$ diferenciando a correspondente resposta ao degrau unitário, uma vez que a função impulso unitário é a derivada temporal da função degrau unitário. Uma família de curvas de resposta a impulso unitário dadas pelas Eqs. (6.31) e (6.32) com vários valores de ζ é vista na Fig. 6.20. As curvas $c(t)/\omega_n$ são mostradas em função da variável adimensional $\omega_n t$, e portanto são funções apenas de ζ. Para os casos de amortecimento crítico e sobreamortecimento, a resposta ao impulso unitário é sempre positiva ou nula; *i.e.*, $c(t) \geq 0$. Isto pode ser visto nas Eqs. (6.32) e (6.33). Para o caso subamortecido, a resposta ao impulso unitário $c(t)$ oscila em torno de zero e assume valores positivos e negativos.

Da análise anterior podemos concluir que, se a resposta ao impulso unitário $c(t)$ não muda de sinal, o sistema ou é criticamente amortecido ou é sobreamortecido, e neste caso a correspondente resposta ao degrau não tem sobre-sinal mas cresce ou decresce monotonicamente e tende a um valor constante.

O sobre-sinal máximo para a resposta ao impulso unitário do sistema subamortecido ocorre em

$$t = \frac{\tan^{-1}\dfrac{\sqrt{1-\zeta^2}}{\zeta}}{\omega_n\sqrt{1-\zeta^2}} \qquad (0 < \zeta < 1)$$

e o sobre-sinal máximo é

$$c(t)_{max} = \omega_n \exp\left(-\frac{\zeta}{\sqrt{1-\zeta^2}} \tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}\right) \qquad (0 < \zeta < 1)$$

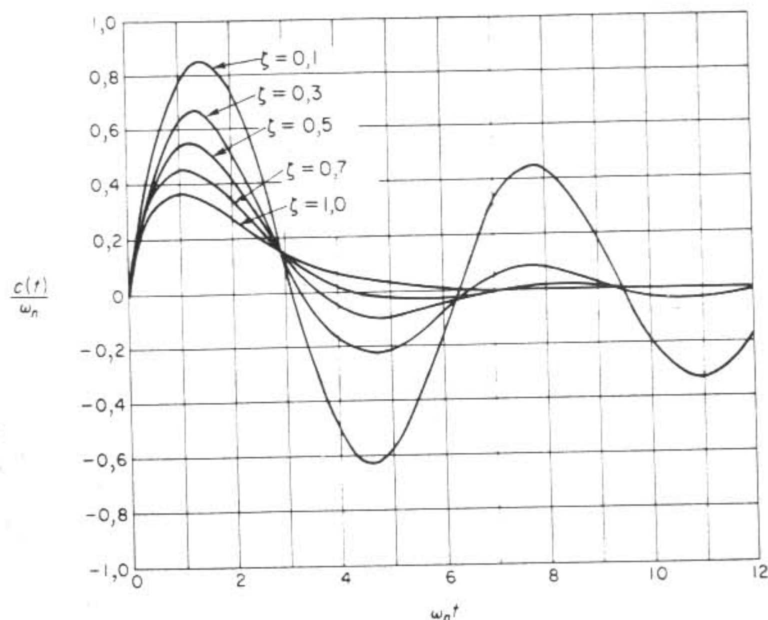


Fig. 6.20 Curvas de resposta a impulso unitário do sistema visto na Fig. 6.10.

Como a função de resposta ao impulso é a derivada temporal da função de resposta degrau unitário, o sobre-sinal máximo M_p para a resposta a degrau unitário pode ser determinado da correspondente resposta ao impulso unitário. Em outras palavras, a área sob a curva de resposta ao impulso unitário de $t = 0$ até o tempo do primeiro cruzamento do zero, como visto na Fig. 6.21, é $1 + M_p$, onde M_p é o sobre-sinal máximo (para a resposta a degrau unitário) dado pela Eq. (6.28). O instante do pico t_p (para a resposta ao degrau unitário) dado pela Eq. (6.27) corresponde ao tempo em que a resposta ao impulso unitário cruza pela primeira vez o eixo.

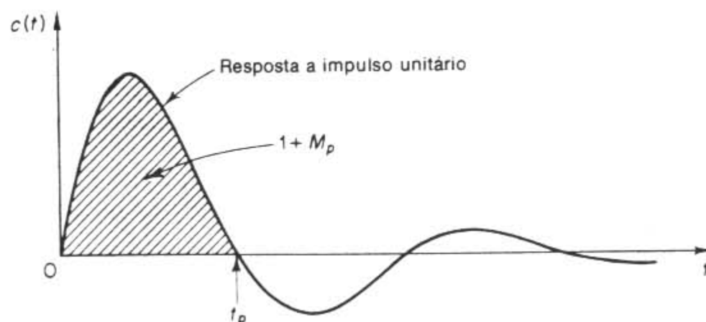


Fig. 6.21 Curva de resposta a impulso unitário do sistema visto na Fig. 6.10.

Erro estacionário para respostas a rampa. A resposta transitória de um sistema de segunda-ordem sujeito a uma entrada em rampa pode ser obtida através de um método simples. (Veja o Problema A.6.5.) Nesta análise, examinaremos o erro estacionário quando o sistema de segunda-ordem é sujeito a tal entrada.

Para o sistema visto na Fig. 6.22, obtemos

$$E(s) = \frac{Js^2 + Fs}{Js^2 + Fs + K} R(s)$$

O erro estacionário para a resposta para a rampa unitária pode ser obtido como segue: Para uma entrada rampa unitária $r(t) = t$, obtemos $R(s) = 1/s^2$. O erro estacionário e_{ss} é

$$\begin{aligned} e_{ss} &= \lim_{s \to 0} sE(s) \\ &= \lim_{s \to 0} s \frac{Js^2 + Fs}{Js^2 + Fs + K} \frac{1}{s^2} \\ &= \frac{F}{K} \\ &= \frac{2\zeta}{\omega_n} \end{aligned}$$

onde

$$\zeta = \frac{F}{2\sqrt{KJ}}, \qquad \omega_n = \sqrt{\frac{K}{J}}$$

Para garantir uma resposta transitória aceitável bem como um erro estacionário admissível em resposta a uma entrada em rampa, ζ não deve ser muito pequeno e ω_n deve ser suficientemente grande. É possível fazer o erro estacionário e_{ss} pequeno fazendo o valor do ganho K grande. (Um valor grande para K tem uma vantagem adicional de suprimir efeitos indesejáveis de zonas mortas, histerese, atrito de Coulomb etc.) Um valor grande para K faria, entretanto, ζ pequeno e aumentaria o sobre-sinal máximo, o que é indesejável.

Portanto é necessário achar uma solução de compromisso entre a valor do erro estacionário para entrada em rampa e o sobre-sinal máximo para uma entrada em degrau unitário. No sistema visto na Fig. 6.22, pode ser difícil obter um compromisso razoável. Nestes casos, então, é desejável se considerar outros tipos de ações de controle que podem melhorar tanto a resposta transitória como o comportamento estacionário. A seguir, consideramos dois métodos para melhorar o com-

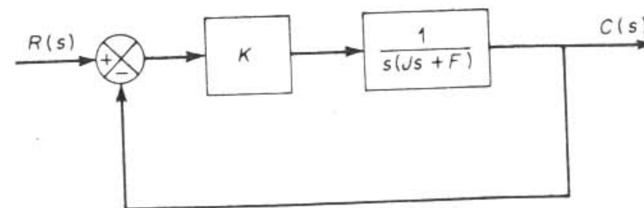


Fig. 6.22 Sistema de controle.

portamento, um usando um controlador proporcional-mais-derivada e o outro usando realimentação de tacômetro.

Controle proporcional-mais-derivada de sistemas de segunda-ordem. Um compromisso entre comportamento aceitável de resposta transitória e de regime estacionário pode ser obtido pelo uso de ações de controle proporcional-mais-derivada.

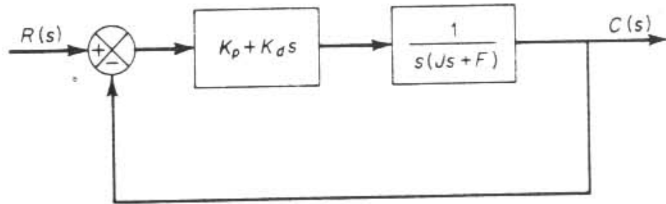


Fig. 6.23 Sistema de controle.

Considere o sistema visto na Fig. 6.23. A função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{K_p + K_d s}{Js^2 + (F + K_d)s + K_p}$$

O erro estacionário para uma entrada em rampa unitária é

$$e_{ss} = \frac{F}{K_p}$$

A equação característica é

$$Js^2 + (F + K_d)s + K_p = 0 \qquad (6.34)$$

O coeficiente de amortecimento efetivo deste sistema é portanto $F + K_d$ ao invés de F. Como o coeficiente de amortecimento deste sistema é

$$\zeta = \frac{F + K_d}{2\sqrt{K_p J}}$$

é possível se fazer tanto o erro estacionário e_{ss} para uma entrada em rampa como o sobre-sinal máximo para uma entrada em degraus pequenos fazendo-se F pequeno, K_p grande, e K_d grande suficiente de tal forma que ζ fique entre 0,4 e 0,7.

A seguir, examinaremos a resposta a degrau unitário do sistema visto na Fig. 6.23. Definamos

$$\omega_n = \sqrt{\frac{K_p}{J}}, \qquad z = \frac{K_p}{K_d}$$

A função de transferência de malha-fechada pode então ser escrita

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{z}\,\frac{s + z}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

Quando um sistema de segunda-ordem tem um zero perto dos pólos de malha-fechada, o comportamento de resposta transitória se torna consideravelmente diferente daquele de um sistema de segunda-ordem sem um zero.

Se o zero em $s = -z$ está localizado perto do eixo $j\omega$, o efeito do zero na resposta a degrau unitário é bastante significativo. Curvas de resposta a degrau típicas deste sistema com $\zeta = 0{,}5$ e com vários valores para $z/(\zeta\omega_n)$ são vistas na Fig. 6.24.

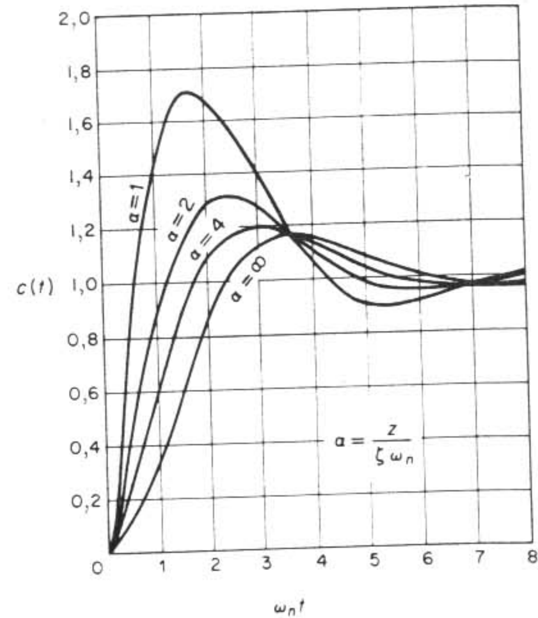


Fig. 6.24 Curvas de resposta a degrau unitário do sistema de segunda-ordem.

$$\frac{C(s)}{R(s)} = \left(\frac{\omega_n^2}{z}\right)\left(\frac{s + z}{s^2 + 2\zeta\omega_n s + \omega_n^2}\right)$$
$$\zeta = 0{,}5$$

Tacômetros. Um outro método para melhorar o desempenho de servomecanismos é adicionar realimentação por tacômetro. Este método é de ajuste fácil e relativamente econômico. Um tacômetro DC é um gerador que produz uma tensão proporcional à sua velocidade de rotação. É usado como um transdutor, convertendo a velocidade do eixo em rotação para uma tensão DC proporcional. Em um tacômetro DC um ímã permanente fornece um fluxo constante no entreferro de ar. Portanto, a tensão induzida pelo tacômetro e pode ser escrita

$$e = K\psi\dot{\theta} = K_1\dot{\theta}$$

onde K é uma constante, ψ é o fluxo no entreferro de ar e $\dot{\theta}$ é a velocidade de rotação. Como ψ é uma constante, $K_1 = K\psi$ também é uma constante. A função de transferência do tacômetro DC é

$$\frac{E(s)}{\Theta(s)} = K_1 s$$

Um diagrama esquemático de um tacômetro DC é visto na Fig. 6.25(a). Embora sua saída seja uma tensão DC, se esta é convertida em uma tensão AC, então o tacômetro DC pode também ser usado em servomecanismos AC.

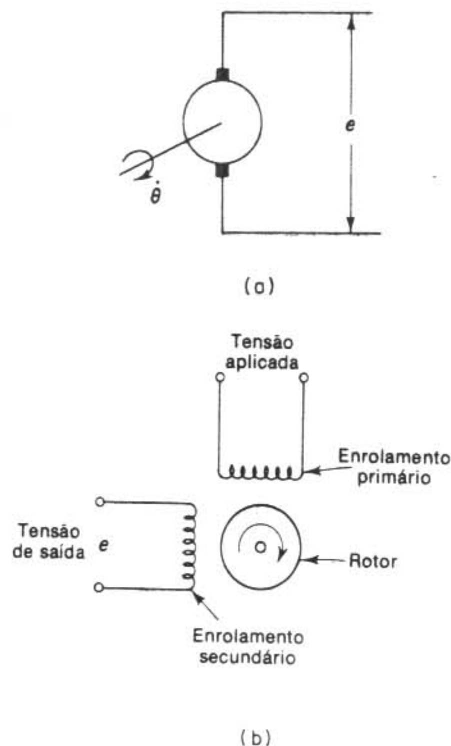


Fig. 6.25 (a) Diagrama esquemático de um tacômetro DC; (b) diagrama esquemático de um tacômetro AC.

O tacômetro AC é um dispositivo similar a um motor de indução de duas fases. (Motores de duas fases geralmente se prestam bem para o papel de tacômetros AC.) A Fig. 6.25(b) mostra o diagrama esquemático de um tacômetro AC. Uma tensão AC fixa é aplicada ao enrolamento primário do tacômetro. O enrolamento secundário é colocado a 90 graus em termos espaciais a partir do enrolamento primário. Portanto, quando o eixo do rotor está estacionário, a tensão de saída é nula. Quando o eixo do rotor está em rotação, a tensão de saída no enrolamento secundário é proporcional à velocidade do rotor. A polaridade da tensão de saída é determinada pela direção de rotação. A função de transferência de um tacômetro AC é

$$\frac{E(s)}{\Theta(s)} = Ks$$

onde $E(s)$ é a transformada de Laplace da tensão de saída, $\Theta(s)$ é a transformada de Laplace da posição do rotor, e K é uma constante. Embora a saída de um tacômetro AC seja uma tensão AC, este tacômetro pode ser usado em um servomecanismo DC se a tensão AC de saída é convertida em uma tensão DC através de um demodulador.

Note que um tacômetro utilizado com finalidades de amortecimento é geralmente construído como parte integral do servomotor.

Servomecanismos com realimentação de velocidade. A derivada do sinal de saída pode ser usada para melhorar o desempenho do sistema. Para se obter a derivada do sinal de posição de saída, é desejável usar um tacômetro ao invés de diferenciar fisicamente o sinal de saída. (Note que a diferenciação amplifica efeitos de ruído. De fato, se ruídos descontínuos estão presentes, a diferenciação amplifica os ruídos descontínuos mais do que sinal útil. Por exemplo, a saída de um potenciômetro é um sinal de tensão descontínua porque, à medida que o terminal móvel do potenciômetro passa nos enrolamentos, tensões são induzidas nos enrolamentos seguintes gerando transitórios. A saída do potenciômetro não deve portanto ser seguida de um elemento de diferenciação.)

Considere o servomecanismo visto na Fig. 6.26(a). Neste dispositivo, o sinal de velocidade, junto com o sinal de posição, é realimentado para a entrada para produzir o sinal de erro atuante. Em qualquer servomecanismo, tal sinal de veloci-

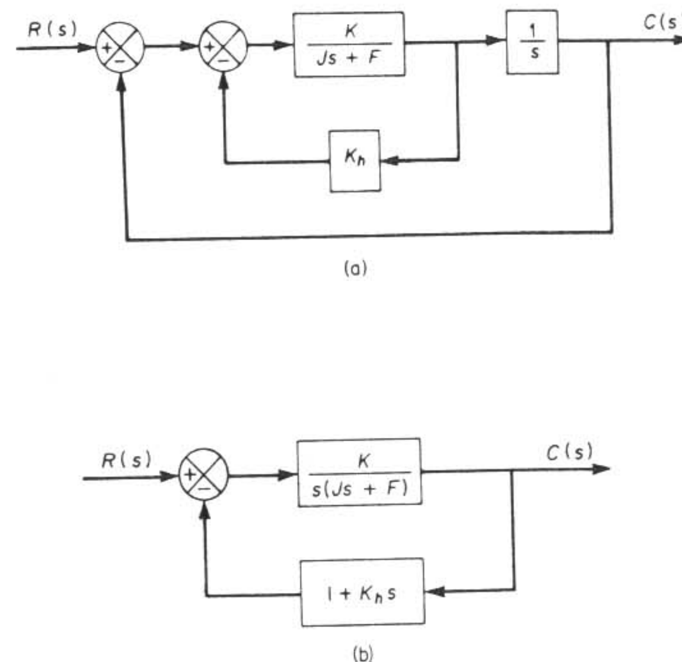


Fig. 6.26 (a) Diagrama de blocos de um servomecanismo; (b) diagrama de blocos simplificado.

dade pode ser facilmente gerado por um tacômetro. O diagrama de blocos visto na Fig. 6.26(a) pode ser simplificado, como visto na Fig. 6.26(b), resultando

$$\frac{C(s)}{R(s)} = \frac{K}{Js^2 + (F + KK_h)s + K}$$

A equação característica é

$$Js^2 + (F + KK_h)s - K = 0 \qquad (6.35)$$

Comparando a Eq. (6.35) com a Eq. (6.34), vemos que elas são da mesma forma. Se KK_h e K fossem iguais a K_d e K_p, respectivamente, então as duas equações seriam idênticas. Portanto podemos esperar que a realimentação de velocidade forneça melhoras semelhantes no desempenho do sistema quando comparada com o uso de ações de controle proporcional-mais-derivada.

O erro estacionário para uma entrada em rampa é

$$e_{ss} = \frac{F}{K}$$

O coeficiente de amortecimento ζ é

$$\zeta = \frac{F + KK_h}{2\sqrt{KJ}} \qquad (6.36)$$

A freqüência natural não amortecida $\omega_n = \sqrt{K/J}$ não é afetada pela realimentação de velocidade. Observando que o sobre-sinal máximo para uma entrada degrau unitário pode ser controlado através do valor do coeficiente de amortecimento ζ, podemos reduzir tanto o erro estacionário e_{ss} quanto o sobre-sinal máximo fazendo F pequeno e K grande e então ajustando a constante de realimentação de velocidade K_h de tal forma que ζ esteja entre 0,4 e 0,7.

Lembrar que a realimentação de velocidade tem o efeito de aumentar o coeficiente de amortecimento sem afetar a freqüência natural não amortecida do sistema.

Exemplo 6.2 Para o sistema visto na Fig. 6.27, determine os valores do ganho K e a constante de realimentação de velocidade K_h de tal forma que o sobre-sinal máximo na resposta a degrau unitário seja 0,2 e o instante do pico seja 1 s. Com estes valores de K e K_h, obtenha o tempo de subida e o tempo de acomodação.

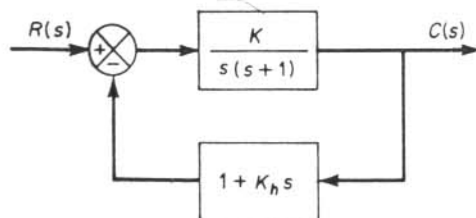


Fig. 6.27 Sistema de controle.

Determinação dos valores de K e K_h: O sobre-sinal máximo M_p é dado pela Eq. (6.28) como sendo

$$M_p = e^{-(\zeta/\sqrt{1-\zeta^2})\pi}$$

Este valor deve ser 0,2. Portanto,

$$e^{-(\zeta/\sqrt{1-\zeta^2})\pi} = 0{,}2$$

ou

$$\frac{\zeta\pi}{\sqrt{1-\zeta^2}} = 1{,}61$$

que fornece

$$\zeta = 0{,}456$$

O instante do pico t_p é especificado como 1 s; portanto, da Eq. (6.27),

$$t_p = \frac{\pi}{\omega_d} = 1$$

ou

$$\omega_d = 3{,}14$$

Como ζ é 0,456, ω_n é

$$\omega_n = \frac{\omega_d}{\sqrt{1-\zeta^2}} = 3{,}53$$

Como a freqüência natural ω_n é igual a $\sqrt{K}$ neste exemplo,

$$K = \omega_n^2 = 12{,}5$$

Então, K_h é, da Eq. (6.36),

$$K_h = \frac{2\sqrt{K}\zeta - 1}{K} = 0{,}178$$

Tempo de subida t_r: Da Eq. (6.26), o tempo de subida t_r é

$$t_r = \frac{\pi - \beta}{\omega_d}$$

onde

$$\beta = \tan^{-1}\frac{\omega_d}{\sigma} = \tan^{-1} 1{,}95 = 1{,}10$$

Portanto, t_r é

$$t_r = 0{,}65 \text{ s}$$

Tempo de acomodação t_s: Para o critério de 2%,

$$t_s = \frac{4}{\sigma} = 2{,}48 \text{ s}$$

Para o critério de 5%,

$$t_s = \frac{3}{\sigma} = 1{,}86 \text{ s}$$

6.5 SISTEMAS DE ORDEM SUPERIOR

Nesta seção, discutiremos inicialmente a resposta a degrau unitário de um tipo particular de sistema de terceira-ordem. Em seguida apresentaremos, em termos gerais, uma análise de resposta transitória de sistemas de ordem superior. Finalmente, discutiremos análise de estabilidade no plano complexo.

Resposta a degrau unitário de sistemas de terceira-ordem. Discutiremos a resposta a degrau unitário de um sistema de terceira-ordem freqüentemente encontrado, cuja função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2 p}{(s^2 + 2\zeta\omega_n s + \omega_n^2)(s + p)}$$

A resposta a degrau unitário deste sistema pode ser obtida como segue:

$$c(t) = 1 - \frac{e^{-\zeta\omega_n t}}{\beta\zeta^2(\beta - 2) + 1}\left\{\beta\zeta^2(\beta - 2)\cos\sqrt{1-\zeta^2}\,\omega_n t + \frac{\beta\zeta[\zeta^2(\beta - 2) + 1]}{\sqrt{1-\zeta^2}}\text{sen}\sqrt{1-\zeta^2}\,\omega_n t\right\} - \frac{e^{-pt}}{\beta\zeta^2(\beta - 2) + 1} \qquad (t \geq 0)$$

onde

$$\beta = \frac{p}{\zeta\omega_n}$$

Note que, como

$$\beta\zeta^2(\beta - 2) + 1 = \zeta^2(\beta - 1)^2 + (1 - \zeta^2) > 0$$

o coeficiente do termo e^{-pt} é sempre negativo.

O efeito do pólo real $s = -p$ na resposta a degrau unitário é de reduzir o sobre-sinal máximo e aumentar o tempo de acomodação. A Fig. 6.28 mostra curvas de resposta a degrau unitário deste sistema de terceira-ordem com $\zeta = 0{,}5$. O quociente $\beta = p/(\zeta\omega_n)$ é um parâmetro na família de curvas.

Se o pólo real está localizado à direita dos pólos complexos-conjugados, então há uma tendência para resposta lenta. O sistema se comportará como um sistema sobreamortecido. Os pólos complexos-conjugados adicionarão ondulação na curva de resposta.

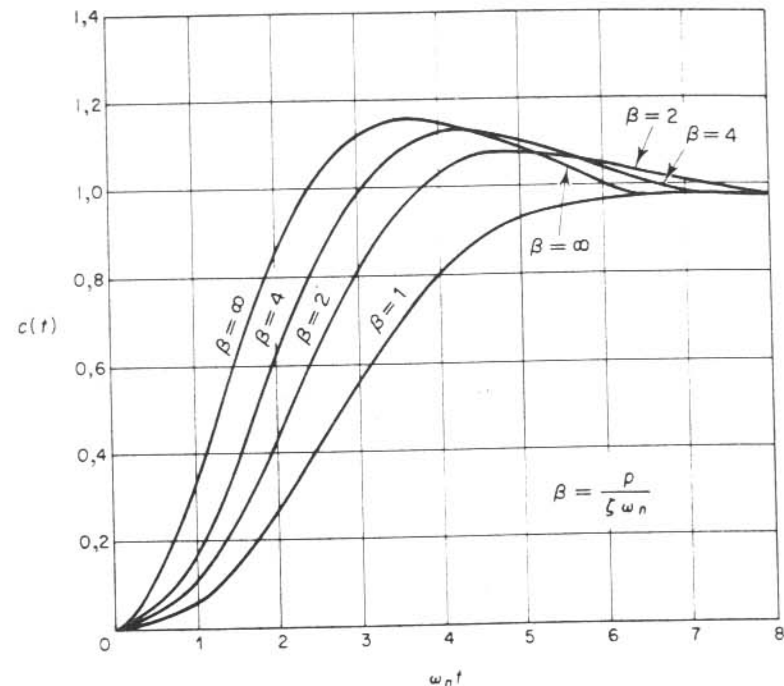


Fig. 6.28 Curvas de resposta a degrau unitário de um sistema de terceira-ordem.

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2 p}{(s^2 + 2\zeta\omega_n s + \omega_n^2)(s + p)}, \qquad \zeta = 0{,}5$$

Resposta transitória de sistemas de ordem superior. Considere o sistema visto na Fig. 6.29. A função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} \qquad (6.37)$$

Em geral, $G(s)$ e $H(s)$ são dados como relações de polinômios em s, ou

$$G(s) = \frac{p(s)}{q(s)} \qquad \text{ou} \qquad H(s) = \frac{n(s)}{d(s)}$$

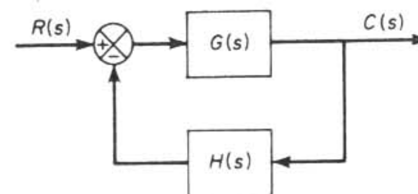


Fig. 6.29 Sistema de controle.

onde $p(s)$, $q(s)$, $n(s)$, e $d(s)$ são polinômios em s. A função de transferência de malha-fechada dada pela Eq. (6.37) pode ser escrita então

$$\frac{C(s)}{R(s)} = \frac{p(s)d(s)}{q(s)d(s) + p(s)n(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \cdots + b_{m-1}s + b_m}{a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}$$

Para determinar a resposta transitória deste sistema para qualquer entrada, é necessário fatorar o polinômio do denominador. (Vários métodos são disponíveis para este fatoramento de polinômios. Neste livro apresentaremos uma técnica conveniente chamada método do lugar das raízes. Veja Cap. 8.) Uma vez que o polinômio do denominador foi fatorado, $C(s)/R(s)$ pode ser escrita

$$\frac{C(s)}{R(s)} = \frac{K(s + z_1)(s + z_2) \cdots (s + z_m)}{(s + p_1)(s + p_2) \cdots (s + p_n)} \tag{6.38}$$

Examinemos o comportamento da resposta deste sistema para uma entrada degrau unitário. Supomos que os pólos de malha-fechada são todos distintos. (Este é geralmente o caso em sistemas práticos.) Para uma entrada degrau unitário, a Eq. (6.38) pode ser escrita

$$C(s) = \frac{a}{s} + \sum_{i=1}^{n} \frac{a_i}{s + p_i} \tag{6.39}$$

onde a_i é o resíduo do pólo em $s = -p_i$.

Se todos os pólos de malha-fechada estão no semiplano esquerdo do plano s, as magnitudes relativas dos resíduos determinam a importância relativa dos componentes na forma expandida de $C(s)$. Se há um zero de malha-fechada perto de um pólo de malha-fechada, então o resíduo neste pólo é pequeno e o coeficiente do termo de resposta transitória correspondendo a este pólo se torna pequeno. Um par de pólos e zeros localizados perto um do outro efetivamente se cancelarão. Se o pólo está localizado muito longe da origem, o resíduo neste pólo pode ser pequeno. Os transitórios correspondendo a tal pólo remoto são pequenos e de curta duração. Os termos na forma expandida de $C(s)$, tendo resíduos muito pequenos, pouco contribuem para a resposta transitória e podem ser desprezados. Se isto é feito, o sistema de ordem superior pode ser aproximado por um de ordem inferior. (Tal aproximação freqüentemente nos permite estimar as características de resposta de um sistema de ordem superior a partir daquelas de um sistema simplificado.)

Os pólos de $C(s)$ consistem em pólos reais e pares de pólos complexos-conjugados. Um par de pólos complexos-conjugados fornece um termo de segunda-ordem em s. Como a forma fatorada da equação característica de ordem superior consiste em termos de primeira e segunda-ordem, então a Eq. (6.39) pode ser reescrita:

$$C(s) = \frac{K \prod_{i=1}^{m} (s + z_i)}{s \prod_{j=1}^{q} (s + p_j) \prod_{k=1}^{r} (s^2 + 2\zeta_k \omega_k s + \omega_k^2)} \tag{6.40}$$

onde $q + 2r = n$. Se os pólos de malha-fechada são distintos, a Eq. (6.40) pode ser expandida em frações parciais como segue:

$$C(s) = \frac{a}{s} + \sum_{j=1}^{q} \frac{a_j}{s + p_j} + \sum_{k=1}^{r} \frac{b_k(s + \zeta_k \omega_k) + c_k \omega_k \sqrt{1 - \zeta_k^2}}{s^2 + 2\zeta_k \omega_k s + \omega_k^2}$$

Desta última equação, podemos ver que a resposta de um sistema de ordem superior é composta de um certo número de termos envolvendo as funções simples vistas nas respostas de sistemas de primeira e segunda-ordem. A resposta a degrau unitário $c(t)$, a transformada inversa de Laplace de $C(s)$, é então

$$c(t) = a + \sum_{j=1}^{q} a_j e^{-p_j t} + \sum_{k=1}^{r} b_k e^{-\zeta_k \omega_k t} \cos \omega_k \sqrt{1 - \zeta_k^2}\, t + \sum_{k=1}^{r} c_k e^{-\zeta_k \omega_k t} \operatorname{sen} \omega_k \sqrt{1 - \zeta_k^2}\, t \qquad (t \geq 0) \tag{6.41}$$

Se todos os pólos de malha-fechada estão no semiplano esquerdo do plano s, então os termos exponenciais e os termos amortecidos exponencialmente na Eq. (6.41) tenderão a zero quanto t cresce. A saída em regime estacionário é então $c(\infty) = a$.

Vamos supor que o sistema em consideração é estável. Então os pólos de malha-fechada que estão localizados longe do eixo $j\omega$ têm partes reais negativas de valor grande. Os termos exponenciais que correspondem a estes pólos decaem rapidamente a zero. (Note que a distância horizontal de um pólo de malha-fechada ao eixo $j\omega$ determina o tempo de acomodação de transitórios devidos àquele pólo. Quanto menor a distância, maior é o tempo de acomodação.)

A curva de resposta de um sistema estável de ordem superior é a soma de um certo número de curvas exponenciais e curvas senoidais amortecidas. Exemplos de curvas de resposta a degrau de sistemas de ordem superior são vistos na Fig. 6.30.

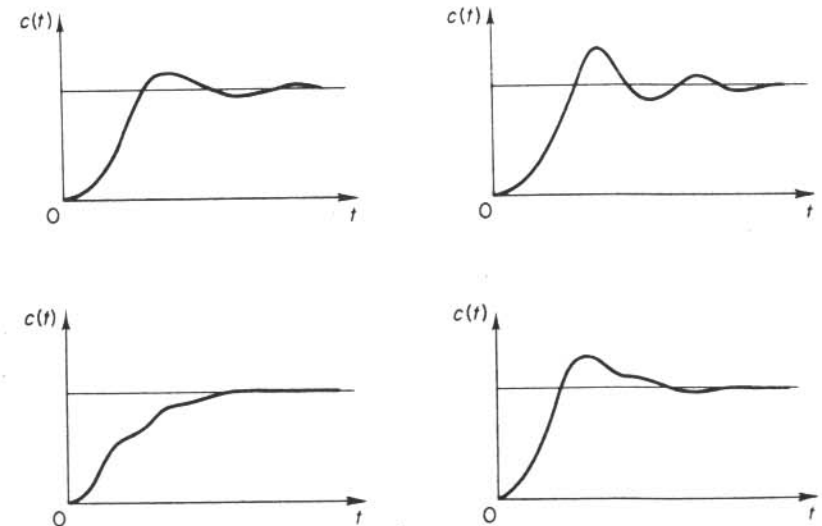


Fig. 6.30 Curvas de resposta a degrau de sistemas de ordem superior.

Uma característica particular de tais curvas de resposta é que pequenas oscilações são superpostas em oscilações maiores ou sobre curvas exponenciais. Componentes de decaimento rápido têm significância somente na parte inicial da resposta transitória.

Lembre-se de que o tipo de resposta transitória é determinado pelos pólos de malha-fechada, ao passo que a forma da resposta transitória é determinada principalmente pelos zeros de malha-fechada. Como vimos antes, os pólos da entrada $R(s)$ fornecem os termos de resposta estacionária na solução, enquanto que os pólos de $C(s)/R(s)$ entram nos termos de resposta transitória exponenciais. (Os pólos de malha-fechada sempre aparecem na solução temporal como termos de resposta exponenciais.) Os zeros de $C(s)/R(s)$ não afetam os expoentes nos termos de resposta exponencial, mas sim as magnitudes e sinais dos resíduos. (A curva de resposta temporal depende dos coeficientes dos termos individuais exponenciais, bem como dos seus expoentes.)

Pólos de malha-fechada dominantes. A dominância relativa de pólos de malha-fechada é determinada pela relação das partes reais dos pólos de malha-fehcada, bem como pelos valores relativos dos resíduos calculados nos pólos de malha-fechada. Os valores dos resíduos dependem tanto dos pólos como dos zeros de malha-fechada.

Se a relação das partes reais excede cinco, e não há zeros na vizinhança, então os pólos de malha-fechada mais perto do eixo $j\omega$ dominarão a resposta transitória porque estes pólos correspondem a termos de resposta transitória que decaem lentamente. Os pólos de malha-fechada que têm efeitos dominantes sobre o comportamento de resposta transitória são chamados pólos *dominantes*. Freqüentemente, os pólos de malha-fechada dominantes ocorrem na forma de pares complexos-conjugados. Os pólos de malha-fechada dominantes são os mais importantes entre os pólos de malha-fechada.

O ganho de um sistema de ordem superior é freqüentemente ajustado até que exista um par de pólos de malha-fechada complexos-conjugados dominantes. A presença de tais pólos em um sistema estável reduz o efeito de tais não-linearidades como zona-morta, histerese e atrito de Coulomb.

Lembre-se que, embora o conceito de pólos de malha-fechada dominantes seja útil para estimar o comportamento dinâmico de um sistema de malha-fechada, devemos ser cuidadosos e verificar se as hipóteses são satisfeitas antes de utilizar este conceito.

Resposta não-oscilatória de sistemas de ordem superior. Se o sistema de malha-fechada não tem pólos complexos-conjugados, então a resposta transitória é não-oscilatória. Transitórios não-oscilatórios podem ser colocados em gráfico semilog, e se a maior constante de tempo for dominante, o gráfico tenderá a uma linha reta cuja inclinação é determinada pela constante de tempo. A subtração da parte da resposta transitória devida a esta constante de tempo fornecerá outras constantes de tempo menores.

Análise de estabilidade no plano complexo. A estabilidade de um sistema linear a malha-fechada pode ser determinada pela localização dos pólos de malha-fechada no plano s. Se qualquer destes pólos está no semiplano direito do plano s, então, para tempos crescentes eles vão resultar no modo dominante, e a resposta transitória aumenta monotonicamente ou oscila com amplitude crescente. Isto representa um sistema instável. Para tal sistema, tão logo a fonte de energia seja ligada, a saída poderá crescer com o tempo. Se não existir saturação no sistema ou algum obstáculo ou breque mecânico, então o sistema pode vir a sofrer danificação e parar de

funcionar uma vez que a resposta de um sistema físico real não pode aumentar indefinidamente. Portanto, pólos de malha-fechada no semiplano direito não são permissíveis no sistema de controle linear usual. Se todos os pólos de malha-fechada estão no semiplano esquerdo do plano s, qualquer resposta transitória alcança um equilíbrio. Isto representa um sistema estável.

O fato de um sistema linear ser estável ou instável é uma propriedade do sistema em si e não depende da entrada do sistema. Os pólos da entrada, ou função excitadora, não afetam a propriedade de estabilidade do sistema, mas contribuem somente para os termos de resposta estacionária na solução. Portanto, o problema de estabilidade absoluta pode ser resolvido de imediato não escolhendo pólos de malha-fechada no semiplano direito do plano s, incluindo o eixo $j\omega$. (Matematicamente, pólos de malha-fechada no eixo $j\omega$ resultarão em oscilações, com amplitude nem decrescente nem crescente com o tempo. Em casos práticos, onde ruído se faz presente, a amplitude das oscilações pode crescer com uma taxa determinada pela potência do ruído. Portanto, um sistema de controle não deve ter pólos de malha-fechada no eixo $j\omega$.)

Note que o mero fato de que todos os pólos de malha-fechada estão no semiplano esquerdo do plano s não garante características satisfatórias para a resposta transitória. Se há pólos de malha-fechada complexos-conjugados dominantes perto do eixo $j\omega$, a resposta transitória pode apresentar oscilações excessivas ou ser muito lenta. Portanto, para garantir uma característica de resposta transitória rápida e ainda bem amortecida, é necessário que os pólos de malha-fechada do sistema se localizem em regiões particulares do plano complexo, como por exemplo, a região limitada pela área hachurada na Fig. 6.31.

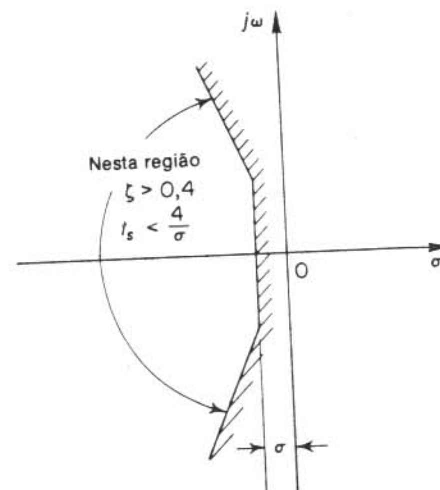


Fig. 6.31 Região do plano complexo satisfazendo as condições $\zeta > 0{,}4$ e $t_s < 4/\sigma$.

Uma vez que estabilidade relativa e desempenho transitório de um sistema de controle de malha-fechada estão diretamente relacionados com a configuração de pólos e zeros de malha-fechada no plano s, freqüentemente se torna necessário ajustar um ou mais parâmetros do sistema para obter configurações satisfatórias. Os efeitos de variação de parâmetros do sistema nos pólos de malha-fechada serão discutidos com detalhes no Cap. 8.

6.6 CRITÉRIO DE ESTABILIDADE DE ROUTH

O problema mais importante em sistemas de controle linear é o da estabilidade. Ou seja, em que condições um sistema se tornará instável? Se ele é instável, como devemos estabilizá-lo? Na Seção 6.5 foi dito que um sistema de controle é estável se e somente se todos os pólos de malha-fechada estão no semiplano esquerdo do plano s. Como a maioria dos sistemas de malha fechada lineares tem funções de transferência de malha-fechada da forma

$$\frac{C(s)}{R(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \cdots + b_{m-1} s + b_m}{a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n} = \frac{B(s)}{A(s)}$$

onde os a's e b's são constantes e $m \leq n$, devemos inicialmente fatorar os polinômios $A(s)$ para achar os pólos de malha-fechada. Este processo leva muito tempo para polinômios de ordem maior que 2. Um critério simples, conhecido como o critério de estabilidade de Routh, nos permite determinar o número de pólos de malha-fechada que estão no semiplano direito s sem ter que fatorar o polinômio.

Critério de estabilidade de Routh. O critério de estabilidade de Routh nos diz se há ou não raízes positivas de uma equação polinomial sem ter que resolver a equação. Este critério de estabilidade se aplica a polinômios com apenas um número finito de termos. Quando o critério é aplicado a um sistema de controle, pode-se obter informação sobre estabilidade absoluta diretamente dos coeficientes da equação característica.

O procedimento no critério de estabilidade de Routh é o seguinte:

1. Escreva o polinômio em s na seguinte forma:

$$a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n = 0 \tag{6.42}$$

onde os coeficientes são grandezas reais. Supomos que $a_n \neq 0$; isto é, qualquer raiz nula foi removida.

2. Se qualquer dos coeficientes é zero ou negativo na presença de pelo menos um coeficiente positivo, então há uma raiz ou raízes que são imaginárias ou que têm partes reais positivas. Portanto, em um caso como este, o sistema não é estável. Se estivermos interessados apenas na estabilidade absoluta, não há necessidade de seguir com o procedimento. Note que todos os coeficientes devem ser positivos. Esta é uma condição necessária, como pode ser visto pelo seguinte argumento: Um polinômio em s tendo coeficientes reais pode ser sempre fatorado em fatores lineares e quadráticos, tais como $(s + a)$ e $(s^2 + bs + c)$, onde a, b e c são reais. Os fatores lineares resultam em raízes reais do polinômio, e os fatores quadráticos, em raízes complexas. O fator $(s^2 + bs + c)$ resulta em raízes com partes reais negativas somente se b e c são ambos positivos. Para que todas as raízes tenham partes reais negativas, as constantes a, b, c etc. em todos os fatores devem ser positivas. O produto de qualquer número de fatores lineares e quadráticos contendo apenas coeficientes positivos sempre resulta em um polinômio com coeficientes positivos. É importante notar que a condição de que todos os coeficientes sejam positivos não é suficiente para assegurar estabilidade. A condição necessária mas não suficiente para estabilidade é que os coeficientes da Eq. (6.42) estejam todos presentes e todos tenham um sinal positivo. (Se todos os a's são negativos, eles podem ser tornados positivos multiplicando-se ambos os lados da equação por -1.)

3. Se todos os coeficientes são positivos, arranje os coeficientes do polinômio em linhas e colunas de acordo com o seguinte padrão:

$$\begin{array}{lllllll}
s^n & a_0 & a_2 & a_4 & a_6 & \cdot & \cdot \; \cdot \\
s^{n-1} & a_1 & a_3 & a_5 & a_7 & \cdot & \cdot \; \cdot \\
s^{n-2} & b_1 & b_2 & b_3 & b_4 & \cdot & \cdot \; \cdot \\
s^{n-3} & c_1 & c_2 & c_3 & c_4 & \cdot & \cdot \; \cdot \\
s^{n-4} & d_1 & d_2 & d_3 & d_4 & \cdot & \cdot \; \cdot \\
\cdot & \cdot & \cdot & & & & \\
\cdot & \cdot & \cdot & & & & \\
\cdot & \cdot & \cdot & & & & \\
s^2 & e_1 & e_2 & & & & \\
s^1 & f_1 & & & & & \\
s^0 & g_1 & & & & &
\end{array}$$

Os coeficientes b_1, b_2, b_3 etc. são calculados como segue:

$$b_1 = \frac{a_1 a_2 - a_0 a_3}{a_1}$$

$$b_2 = \frac{a_1 a_4 - a_0 a_5}{a_1}$$

$$b_3 = \frac{a_1 a_6 - a_0 a_7}{a_1}$$

$$\cdot \quad \cdot \quad \cdot$$

O cálculo dos b's continua até que os restantes sejam todos nulos. O mesmo padrão de multiplicação em cruz dos coeficientes das duas linhas anteriores é seguido para calcular os c's d's, e's etc. Ou seja,

$$c_1 = \frac{b_1 a_3 - a_1 b_2}{b_1}$$

$$c_2 = \frac{b_1 a_5 - a_1 b_3}{b_1}$$

$$c_3 = \frac{b_1 a_7 - a_1 b_4}{b_1}$$

$$\cdot \quad \cdot \quad \cdot$$

e

$$d_1 = \frac{c_1 b_2 - b_1 c_2}{c_1}$$

$$d_2 = \frac{c_1 b_3 - b_1 c_3}{c_1}$$

$$\cdot \quad \cdot \quad \cdot$$

Este processo continua até que a linha n-ésima tenha sido completada. A tabela completa de coeficientes é triangular. Note que, ao desenvolver a tabela, uma linha

inteira pode ser dividida ou multiplicada por um número positivo visando simplificar os cálculos subseqüentes sem alterar a conclusão de estabilidade.

O critério de estabilidade de Routh diz que o número de raízes da Eq. (6.42) com partes reais positivas é igual ao número de mudanças de sinal dos coeficientes da primeira coluna da tabela. Deve-se notar que o valor exato dos termos da primeira coluna não é de importância; apenas os sinais são necessários. A condição necessária e suficiente para que todas as raízes da Eq. (6.42) fiquem no semiplano esquerdo do plano s é que todos os coeficientes da Eq. (6.42) sejam positivos e que todos os termos da primeira coluna da tabela tenham sinal positivo.

Exemplo 6.3 Vamos aplicar o critério de estabilidade de Routh para o seguinte polinômio de terceira-ordem:

$$a_0s^3 + a_1s^2 + a_2s + a_3 = 0$$

onde todos os coeficientes são números positivos. A tabela de coeficientes se torna

s^3	a_0	a_2
s^2	a_1	a_3
s^1	$\dfrac{a_1a_2 - a_0a_3}{a_1}$	
s^0	a_3	

A condição de que todas as raízes tenham partes reais negativas é dada por

$$a_1a_2 > a_0a_3$$

Exemplo 6.4 Considere o seguinte polinômio:

$$s^4 + 2s^3 + 3s^2 + 4s + 5 = 0$$

Vamos seguir o procedimento apresentado e construir a tabela de coeficientes. (As duas primeiras linhas podem ser obtidas diretamente do polinômio dado. Os demais termos são obtidos a partir destes. Se qualquer dos coeficientes é inexistente, este pode ser substituído por um zero na tabela.)

s^4	1	3	5
s^3	2	4	0
s^2	1	5	
s^1	-6		
s^0	5		

s^4	1	3	5	
s^3	~~2~~	~~4~~	~~0~~	A segunda linha é dividida por dois
	1	2	0	
s^2	1	5		
s^1	-3			
s^0	5			

Neste exemplo, o número de mudanças de sinal dos coeficientes da primeira coluna é igual a dois. Isto significa que há duas raízes com partes reais positivas. Note que o resultado não se altera quando os coeficientes de qualquer linha são multiplicados ou divididos por um número positivo para simplificar os cálculos.

Casos especiais. Se um termo da primeira coluna é nulo, mas os termos restantes não são nulos ou não há termo restante, então o termo nulo é substituído por um número positivo muito pequeno ϵ e o resto da tabela é calculado. Por exemplo, considere a seguinte equação:

$$s^3 + 2s^2 + s + 2 = 0 \qquad (6.43)$$

A tabela de coeficientes é

s^3	1	1
s^2	2	2
s^1	$0 \approx \epsilon$	
s^0	2	

Se o sinal do coeficiente acima do zero (ϵ) é o mesmo que o sinal do coeficiente abaixo, isto indica que há um par de raízes imaginárias. De fato, a Eq. (6.43) tem duas raízes em $s = \pm j$.

Se, entretanto, o sinal do coeficiente acima do zero (ϵ) é oposto ao do coeficiente abaixo, isto indica que há uma mudança de sinal.* Por exemplo, para a seguinte equação:

$$s^3 - 3s + 2 = (s - 1)^2(s + 2) = 0$$

a tabela de coeficientes é

	s^3	1	-3
uma mudança de sinal:	s^2	$0 \approx \epsilon$	2
	s^1	$-3 - \dfrac{2}{\epsilon}$	
uma mudança de sinal:	s^0	2	

Há duas mudanças de sinal dos coeficientes da primeira coluna. Isto está de acordo com o resultado correto indicado pela forma fatorada da equação polinomial.

Se todos os coeficientes em uma linha calculada forem nulos, isto indica que há raízes de igual módulo mas radialmente opostas no plano s, isto é, duas raízes reais com módulo igual mas sinal oposto e/ou duas raízes imaginárias conjugadas. Em tal caso, o cálculo do resto da tabela pode ser continuado formando um polinômio auxiliar com os coeficientes da última linha e usando os coeficientes da derivada deste polinômio na linha seguinte. Tais raízes com igual módulo e situadas radialmente opostas no plano s podem ser determinadas resolvendo o polinômio auxiliar, que é sempre par. Para um polinômio auxiliar de grau $2n$, há n pares de raízes iguais e opostas. Por exemplo, considere a seguinte equação:

$$s^5 + 2s^4 + 24s^3 + 48s^2 - 25s - 50 = 0$$

---

*N. do T.: Este caso de coeficiente nulo na primeira coluna não é tão simples como descrito pelo autor, podendo em casos mais complexos levar a erro, como descrito em On the Routh-Hurwitz Criterion, T. S. Chang e C. T. Chen; *IEEE Trans. Automatic Control*, 19, pp. 250-251, 1974, e nos comentários publicados na mesma revista, vol. 20, pp. 295-297, 1975.

A tabela de coeficiente é

s^5	1	24	−25	
s^4	2	48	−50	← Polinômio auxiliar $P(s)$
s^3	0	0		

Os termos na linha s^3 são todos nulos. O polinômio auxiliar é então formado dos coeficientes da linha s^4. O polinômio auxiliar $P(s)$ é

$$P(s) = 2s^4 + 48s^2 - 50$$

que indica que há dois pares de raízes de igual módulo e sinal oposto. Estes pares são obtidos resolvendo-se a equação polinomial auxiliar $P(s) = 0$. A derivada de $P(s)$ com respeito a s é

$$\frac{dP(s)}{ds} = 8s^3 + 96s$$

Os termos na linha s^3 são substituídos pelos coeficientes da última equação, ou seja, 8 e 96. A tabela de coeficientes, então, se torna

s^5	1	24	−25	
s^4	2	48	−50	
s^3	8	96		← Coeficientes de $dP(s)/ds$
s^2	24	−50		
s^1	112,7	0		
s^0	−50			

Vemos que há uma mudança de sinal na primeira coluna da nova tabela. Portanto, a equação original tem uma raiz com parte real positiva. Resolvendo a equação polinomial auxiliar, obtemos as raízes

$$2s^4 + 48s^2 - 50 = 0$$

obtemos

$$s^2 = 1, \qquad s^2 = -25$$

ou

$$s = \pm 1, \qquad s = \pm j5$$

Estes dois pares de raízes são uma parte das raízes da equação original. De fato, a equação original pode ser escrita em forma fatorada como segue:

$$(s + 1)(s - 1)(s + j5)(s - j5)(s + 2) = 0$$

Claramente, a equação original tem uma raiz com parte real positiva.

Análise de estabilidade relativa. O critério de estabilidade de Routh provê a resposta à questão de estabilidade absoluta. Isto, em muitos casos práticos, não é suficiente. Normalmente requeremos informação sobre a estabilidade relativa do sistema. Um método útil de se examinar a estabilidade relativa é o de deslocar o eixo do plano s e aplicar o critério de estabilidade de Routh. Ou seja, substituímos

$$s = z - \sigma \ (\sigma = \text{constante})$$

na equação característica do sistema, escrevemos o polinômio em termos de z, e então aplicamos o critério de estabilidade de Routh para o novo polinômio em z. O número de mudanças de sinal na primeira coluna da tabela feita para o polinômio em z é igual ao número de raízes que estão localizadas para a direita da linha vertical $s = -\sigma$. Portanto, este teste revela o número de raízes que estão à direita da linha vertical $s = -\sigma$.

Aplicação do critério de estabilidade de Routh para a análise de sistemas de controle. O critério de estabilidade de Routh é de utilidade limitada na análise de sistemas de controle lineares, principalmente porque ele não sugere como melhorar a estabilidade relativa ou como estabilizar um sistema instável. Entretanto, é possível determinar os efeitos da modificação de um ou dois parâmetros de um sistema examinando os valores que causam instabilidade. Em seguida, consideraremos o problema de determinar a região de estabilidade para um valor de um parâmetro.

Considere o sistema visto na Fig. 6.32. Vamos determinar a região de valores de K para haver estabilidade.

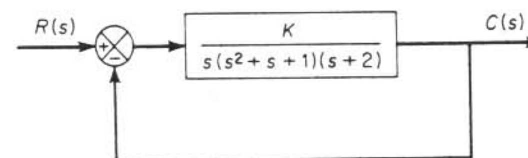


Fig. 6.32 Sistema de controle.

A função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{K}{s(s^2 + s + 1)(s + 2) + K}$$

A equação característica é

$$s^4 + 3s^3 + 3s^2 + 2s + K = 0$$

A tabela dos coeficientes se torna

s^4	1	3	K
s^3	3	2	0
s^2	$\frac{7}{3}$	K	
s^1	$2 - \frac{9}{7}K$		
s^0	K		

Para estabilidade, K deve ser positivo, e todos os coeficientes da primeira coluna devem ser positivos. Portanto,

$$\tfrac{14}{9} > K > 0$$

Quando $K = 14/9$, o sistema se torna oscilatório e, matematicamente, a oscilação é mantida com amplitude constante.

6.7 COMPUTADORES ANALÓGICOS

Esta seção apresenta os princípios de operação de computadores eletrônicos analógicos e as técnicas de montar diagramas computacionais para resolver equações diferenciais e para simular sistemas físicos.

Sabe-se que o computador analógico é uma das ferramentas mais úteis para o engenheiro analisar e projetar tanto sistemas lineares como não-lineares.* Iniciaremos nossa discussão descrevendo os elementos de um computador eletrônico analógico.

Amplificadores operacionais. Um amplificador operacional é um amplificador usado para produzir várias operações matemáticas, como inversão de sinal, soma e integração. É um amplificador DC e tem um ganho muito alto, aproximadamente $10^6 \sim 10^8$. A corrente requerida na entrada de um amplificador operacional é desprezível. A tensão de saída é usualmente limitada a ± 100 volts. A Fig. 6.33(a) mostra um diagrama esquemático de um amplificador operacional. A tensão de saída e_0 e a tensão de entrada e_g são relacionadas pela seguinte equação:

$$e_0 = -Ke_g$$

onde $K = 10^6 \sim 10^8$. Um amplificador operacional, portanto, muda o sinal algébrico.

A Fig. 6.33(b) mostra um amplificador operacional em série com uma impedância de entrada Z_i e com uma impedância Z_0 inserida no elo de realimentação. Como a impedância interna do amplificador é muito alta, a corrente de entrada é essencialmente desprezível, ou

$$i_g \doteqdot 0$$

Então, pela lei de Kirchhoff,

$$i_1 = i_2$$

Tomando a transferência de Laplace de ambos os lados desta equação, obtemos

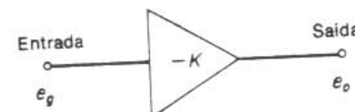
$$I_1(s) = I_2(s)$$

onde

$$I_1(s) = \frac{E_i(s) - E_g(s)}{Z_i(s)}$$

$$I_2(s) = \frac{E_g(s) - E_0(s)}{Z_0(s)}$$

*N. do T.: Com a grande disponibilidade atual de computadores digitais, os computadores analógicos já não são tão importantes, não só em análise como em projeto.



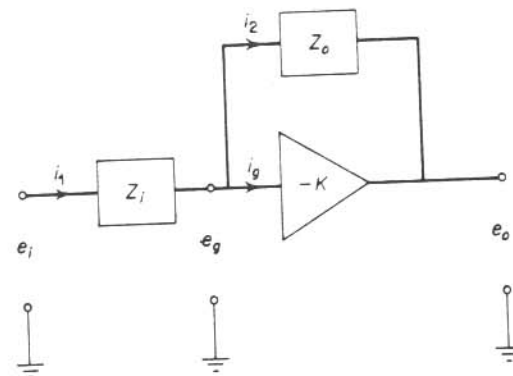


Fig. 6.33 (a) Diagrama esquemático de um amplificador operacional; (b) diagrama esquemático de um amplificador operacional com impedâncias de entrada e de realimentação.

Portanto, obtemos

$$\frac{E_i(s)}{Z_i(s)} = \frac{E_g(s)}{Z_i(s)} + \frac{E_g(s)}{Z_0(s)} - \frac{E_0(s)}{Z_0(s)} \tag{6.44}$$

Observando que

$$E_0(s) = -KE_g(s)$$

e que K é um número muito grande (entre 10^6 e 10^8), obtemos a seguinte equação, desprezando os primeiros dois termos do lado direito da Eq. (6.44):

$$E_0(s) = -\frac{Z_0(s)}{Z_i(s)}E_i(s) \tag{6.45}$$

A Eq. (6.45) é a equação básica relacionando e_i e e_0 do amplificador operacional DC de alto ganho visto na Fig. 6.33(b).

Inversão de sinal. Se usarmos resistores como impedâncias de entrada e de realimentação, como visto na Fig. 6.34, então a Eq. (6.45) se torna

$$E_0(s) = -\frac{R_0}{R_i}E_i(s)$$

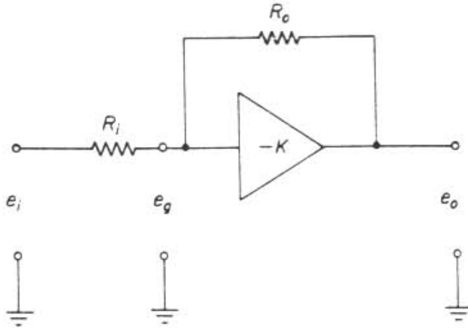
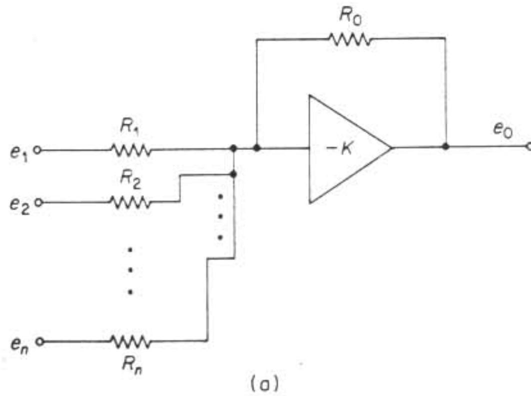


Fig. 6.34 Inversor de sinal.

A tensão de saída é igual à tensão de entrada vezes uma constante $(-R_0/R_i)$, que é negativa. Portanto, a Fig. 6.34 mostra um circuito fornecendo uma inversão de sinal.

Soma. A soma de várias entradas pode ser feita usando-se resistores como impedâncias de entrada e de realimentação do amplificador operacional DC de alto ganho. A Fig. 6.35(a) mostra um diagrama esquemático de um somador que adiciona n entradas. A saída e_0 é

$$e_0 = -\left(\frac{R_0}{R_1}e_1 + \frac{R_0}{R_2}e_2 + \cdots + \frac{R_0}{R_n}e_n\right)$$



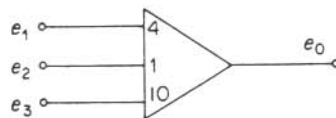


Fig. 6.35 (a) Diagrama esquemático de um somador; (b) símbolo de um somador.

Se, por exemplo, $n = 3$ e $R_0 = 1\ \text{M}\Omega = 10^6\ \Omega$, $R_1 = 0{,}25\ \text{M}\Omega$, $R_2 = 1\ \text{M}\Omega$, e $R_3 = 0{,}1\ \text{M}\Omega$, então,

$$e_0 = -(4e_1 + e_2 + 10e_3)$$

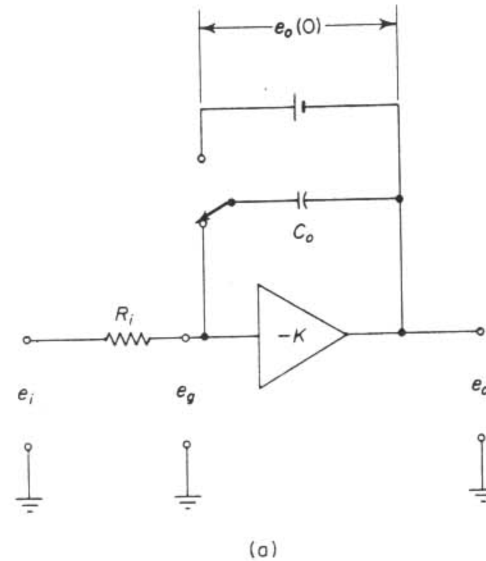
A Fig. 6.35(b) mostra um símbolo usado comumente para tal caso.

Integração. Se um capacitor é usado como a impedância de realimentação e um resistor como impedância de entrada, como visto na Fig. 6.36(a), então a Eq. (6.45) se torna

$$E_0(s) = -\frac{1}{R_iC_0s}E_i(s)$$

ou

$$e_0 = -\frac{1}{R_iC_0}\int e_i\,dt \tag{6.46}$$



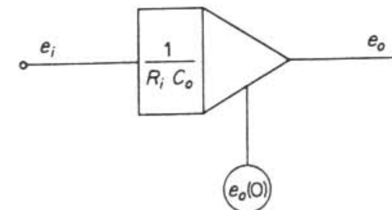


Fig. 6.36 (a) Diagrama esquemático de um integrador; (b) símbolo de um integrador.

Note que ao integrar a entrada e_i para obter a saída e_0, a condição inicial deve ser fornecida. Indicando a condição inicial por $e_0(0)$, podemos escrever a Eq. (6.46) como

$$e_0(t) = -\frac{1}{R_iC_0}\int_0^t e_i(t)\,dt + e_0(0)$$

Cada integrador deve ser inicialmente polarizado por uma tensão DC para dar a condição inicial requerida.

A Fig. 6.36(b) mostra um símbolo comumente usado para um integrador. Se, por exemplo, $R_i = 0{,}25$ MΩ e $C_0 = 1\,\mu$F, então $1/(R_iC_0) = 4$. A condição inicial $e_0(0)$ é indicada no círculo. (Note que alguns amplificadores operacionais usam resistores padrões de 1, 0,25 e 0,10 MΩ e um capacitor padrão de 1 μF. Em tal caso, os valores de $1/R_iC_0$ são iguais somente a 1, 4 ou 10.)

Multiplicação por uma fração. A multiplicação de e_i por uma constante α, onde $0 < \alpha < 1$, pode ser feita com o potenciômetro visto na Fig. 6.37(a). A saída e_0 é

$$e_0 = \frac{R_0}{R_i}e_i$$

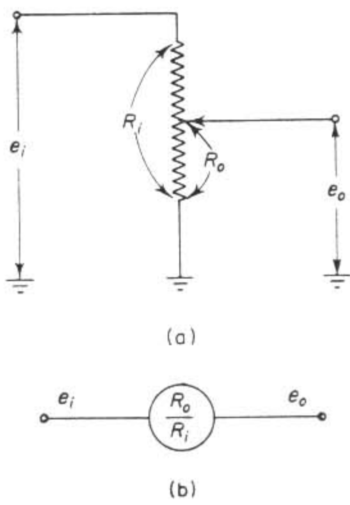


Fig. 6.37 (a) Potenciômetro; (b) símbolo do potenciômetro.

A Fig. 6.37(b) mostra o símbolo comumente usado para um potenciômetro.

Resolvendo equações diferenciais. Vamos ilustrar o procedimento para resolver equações diferenciais pelo uso do computador eletrônico analógico. Considere, como exemplo, a seguinte equação diferencial:

$$\ddot{x} + 10\dot{x} + 16x = 0, \qquad x(0) = 0, \qquad \dot{x}(0) = 8 \tag{6.47}$$

O primeiro passo na montagem de um diagrama de computação é supor que a derivada de ordem mais alta está disponível. Então, resolver a equação diferencial para esta derivada da ordem mais alta. Na equação diferencial do exemplo,

$$\ddot{x} = -10\dot{x} - 16x$$

A variável $\dot{x}$ pode ser obtida integrando $\ddot{x}$, e x pode ser obtida integrando $\dot{x}$. A Fig. 6.38 mostra um diagrama de computação para este sistema. Note que a soma das entradas do primeiro integrador é o termo de derivada mais alta que supomos disponível. Note também que as condições iniciais são indicadas nos círculos.

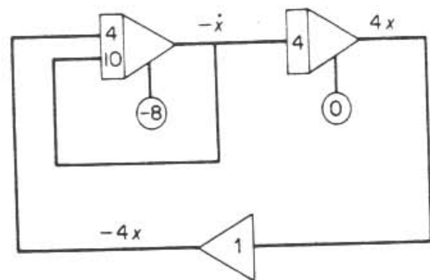


Fig. 6.38 Diagrama de computação analógica para resolver a Eq. (6.47).

É importante lembrar que há uma mudança de sinal associada com cada amplificador operacional. Portanto, se o número de amplificadores operacionais (integradores e somadores) em uma malha-fechada é par, as tensões de saída aumentarão até que saturem. Portanto, o número de amplificadores operacionais em qualquer malha-fechada deve ser ímpar. (No diagrama de computação visto na Fig. 6.38, a malha interna tem um amplificador operacional, e a malha de saída, três.) Isto serve como um teste conveniente para ver se há algum erro na montagem do diagrama de computação.

Geração de funções senoidais. Ilustraremos a seguir a geração de um sinal senoidal, como por exemplo, 10 sen $3t$. Para montar um diagrama de computação analógica, vamos obter uma equação diferencial cuja solução seja 10 sen $3t$.

Seja

$$x(t) = 10\,\text{sen}\,3t$$

Então

$$\ddot{x}(t) = -90\,\text{sen}\,3t$$

Portanto a equação diferencial requerida é

$$\ddot{x} + 9x = 0, \qquad x(0) = 0, \qquad \dot{x}(0) = 30$$

Resolvendo esta equação diferencial para a derivada de ordem mais elevada, obtemos

$$\ddot{x} = -9x$$

Supondo que $\ddot{x}$ é disponível, x pode ser obtido integrando $\ddot{x}$ duas vezes. Um diagrama de computação para este sistema é visto na Fig. 6.39.

Note que a saída do segundo integrador é $9x$ e que a tensão de saída oscila entre ± 90 V. (As condições iniciais determinam a amplitude da oscilação de tensão.)

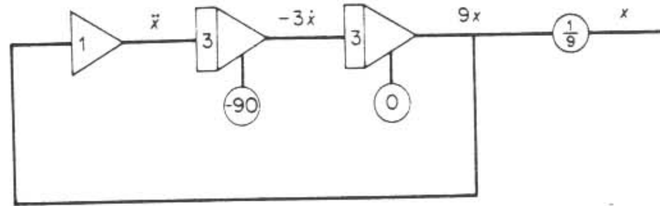


Fig. 6.39 Diagrama de computação analógica para gerar um sinal senoidal $x(t) = 10$ sen $3t$.

Fatores de escala de magnitude. A tensão de saída de qualquer amplificador não deve exceder ± 100 V para evitar saturação. Saturação de voltagem causará erros na solução. Por outro lado, a tensão máxima de qualquer amplificador não deve ser muito pequena. Para assegurar uma precisão adequada, é desejável que a máxima excursão da tensão de saída de qualquer amplificador operacional esteja em torno de ± 70 a ± 90 V.

Ao montar o diagrama de computação, é desejável que a máxima excursão da tensão de saída de qualquer amplificador seja mais ou menos a mesma. A este respeito, a escolha dos fatores de escala de magnitude é de grande importância. Os fatores de escala de magnitude relacionam tensões de saída dos amplificadores com as quantidades físicas correspondentes.

Considere o sistema dado pela Eq. (6.47). Vamos supor que x tem unidades de comprimento em *feet*. Então $x(0) = 0$ ft, e $\dot{x}(0) = 8$ ft/s. Resolvendo a Eq. (6.47) para a derivada de maior ordem, obtemos

$$\ddot{x} = -10\dot{x} - 16x \tag{6.48}$$

Vamos definir k_1, k_2 e k_3 como fatores de escala de magnitude e reescrever a Eq. (6.48) como

$$k_1\ddot{x} = -\frac{10k_1}{k_2}(k_2\dot{x}) - \frac{16k_1}{k_3}(k_3x) \tag{6.49}$$

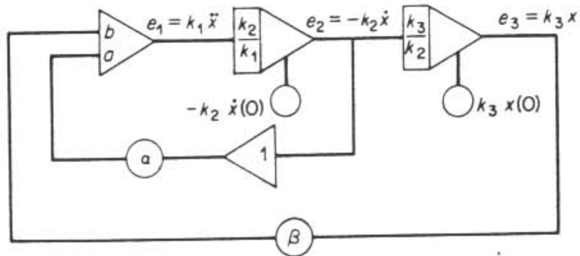


Fig. 6.40 Diagrama de computação analógica sem fatores de escala para resolver a Eq. (6.47).

Referindo-se à Fig. 6.40, a tensão de saída do primeiro amplificador (somador) é $e_1 = k_1\ddot{x}$. A tensão de saída do segundo amplificador (primeiro integrador) é $e_2 = -k_2\dot{x}$, e, finalmente, a tensão de saída do terceiro amplificador (segundo integrador) é $e_3 = k_3x$. Estas tensões de saída devem estar limitadas a ± 100 V.

A freqüência natural não-amortecida deste sistema é 4 rad/s. Uma solução aproximada para $0 < t \ll 1$ é

$$x(t) = 2\,\text{sen}\,4t$$

Portanto, para $0 < t \ll 1$,

$$\dot{x}(t) = 8\cos 4t$$
$$\ddot{x}(t) = -32\,\text{sen}\,4t$$

Note que esta solução só é valida para valores muito pequenos de t. Desta solução, obtemos

Valor máximo de $|x(t)| = 2$
Valor máximo de $|\dot{x}(t)| = 8$
Valor máximo de $|\ddot{x}(t)| = 32$

[Note que em sistemas mais complexos (especialmente em sistemas não-lineares) os valores máximos de $x(t)$, $\dot{x}(t)$, $\ddot{x}(t)$, ... não são disponíveis. Devemos então estimar estes valores máximos. Se os valores estimados estão muito errados, então o diagrama de computação deve ser revisado até que se torne aceitável.] Gostaríamos de escolher k_1, k_2, e k_3 de tal forma que $e_1 = -e_2 = e_3 = 90$ V para os máximos valores de $\ddot{x}$, $\dot{x}$ e x, respectivamente. (Escolhemos 90 V em vez de 100 V como uma estimativa conservadora.) Portanto,

$$k_1 = \frac{90}{|\ddot{x}|_{max}} = \frac{90}{32} = \frac{45}{16}\ \text{volts/ft/s}^2$$

$$k_2 = \frac{90}{|\dot{x}|_{max}} = \frac{90}{8} = \frac{45}{4}\ \text{volts/ft/s}$$

$$k_3 = \frac{90}{|x|_{max}} = \frac{90}{2} = 45\ \text{volts/ft}$$

Portanto

$$\frac{k_2}{k_1} = 4, \qquad \frac{k_3}{k_2} = 4$$

Note que da Fig. 6.40, obtemos

$$k_1\ddot{x} = e_1 = -(-a\alpha e_2 + b\beta e_3) \tag{6.50}$$

Da Eq. (6.49),

$$k_1\ddot{x} = -\frac{10k_1}{k_2}(-e_2) - \frac{16k_1}{k_3}e_3 \tag{6.51}$$

Equacionando as Eqs. (6.50) e (6.51), obtemos

$$a\alpha = \frac{10k_1}{k_2}, \qquad b\beta = \frac{16k_1}{k_3}$$

Portanto,

$$a\alpha = 2{,}5, \qquad b\beta = 1$$

Vamos escolher $a = 4$, $\alpha = 2{,}5/4$, e $b = \beta = 1$. Então, todas as constantes desconhecidas na Fig. 6.40 estão determinadas. O diagrama de computação com fator de escala adequado é visto na Fig. 6.41. As condições iniciais são

$$e_2(0) = -k_2\dot{x}(0) = -\frac{45}{4} \times 8 = -90 \text{ volts}$$

$$e_3(0) = k_3 x(0) = 45 \times 0 = 0 \text{ volts}$$

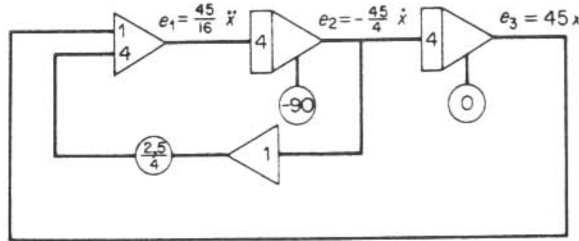


Fig. 6.41 Diagrama de computação analógica com fatores de escala adequados para resolver a Eq. (6.47).

Note que o diagrama visto na Fig. 6.41 não é o diagrama definitivo que gostaríamos de usar como diagrama de computação para resolver a Eq. (6.48). Para minimizar o efeito de ruído e manter a alta precisão, é desejável usar o número mínimo possível de amplificadores. (Em qualquer computador analógico, o número de integradores e somadores é limitado. Ao resolver um problema complexo requerendo muitos integradores e somadores, é necessário usar o número mínimo para cada equação para se economizar componentes.)

No caso do diagrama da Fig. 6.41, é possível eliminar um somador. O diagrama de computação simplificado usando os mesmos fatores de magnitude de escala é visto na Fig. 6.42. Note que como o ganho da malha interna é $4 \times 1 \times (2{,}5/4) \times 4 = 10$, o potenciômetro foi eliminado. Comparando o diagrama de computação visto na Fig. 6.38 com aquele da Fig. 6.42, vemos que o último é melhor do ponto de vista de precisão, uma vez que todos os amplificadores têm excursão sobre uma porção maior de sua região de linearidade.

Fatores de escala de tempo. A escala de tempo relaciona a variável independente do problema físico à variável independente do computador analógico. Para fenômenos que acontecem muito rapidamente, é necessário diminuir a velocidade com que tais problemas são simulados no computador.

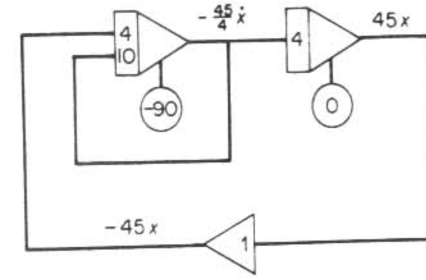


Fig. 6.42 Versão simplificada do diagrama de computação visto na Fig. 6.41.

Façamos com que a equação seguinte relacione o tempo real t em segundos ao tempo do computador (ou tempo de máquina) τ em segundos:

$$\tau = \lambda t$$

onde λ é o fator de escala de tempo. Se λ é escolhido como 0,1, então, 10 s de tempo real são equivalentes a 1 segundo de computador. Isto significa que, se a resposta leva 10 s de tempo real para ser completada, então a resposta é completada em 1 s no computador. Ou vice-versa, se λ é escolhido como sendo 10, então, 1 s de tempo real é equivalente a 10 s de tempo de computador.

Como um exemplo, considere a seguinte equação diferencial:

$$\frac{d^2x}{dt^2} + 10\frac{dx}{dt} + 16x = 20\,\text{sen}\,\omega t, \qquad x(0) = 0, \qquad \left.\frac{dx}{dt}\right|_{t=0} = 8 \tag{6.52}$$

Vamos converter a variável independente de t para τ. Como

$$\tau = \lambda t$$

obtemos

$$\frac{dx}{dt} = \frac{dx}{d\tau}\frac{d\tau}{dt} = \lambda\frac{dx}{d\tau}$$

$$\frac{d^2x}{dt^2} = \lambda^2\frac{d^2x}{d\tau^2}$$

A Eq. (6.52) torna-se então

$$\lambda^2\frac{d^2x}{d\tau^2} + 10\lambda\frac{dx}{d\tau} + 16x = 20\,\text{sen}\frac{\omega}{\lambda}\tau \tag{6.53}$$

Para tornar a solução mais lenta de um fator de 5, substituímos $\lambda = 5$ na Eq. (6.53). A equação para o computador é então

$$25\frac{d^2x}{d\tau^2} + 50\frac{dx}{d\tau} + 16x = 20\,\text{sen}\frac{\omega}{5}\tau$$

As condições iniciais são transformadas para

$$x(0) = 0, \qquad \left.\frac{dx}{d\tau}\right|_{\tau=0} = \frac{1}{\lambda}\left.\frac{dx}{dt}\right|_{t=0} = \frac{1}{5}(8) = 1{,}6$$

Como um outro exemplo, considere a seguinte equação:

$$100\frac{d^2x}{dt^2} + 2\frac{dx}{dt} + 0{,}5x = \operatorname{sen} t, \qquad x(0) = 2, \qquad \left.\frac{dx}{dt}\right|_{t=0} = 5 \tag{6.54}$$

Em virtude da grande diferença nos valores dos coeficientes, a simulação da Eq. (6.54) no computador não é desejável. Se escolhemos para λ um valor tal que os coeficientes de todos os termos são da mesma ordem de magnitude, então, a solução do computador terá em geral menos erro. No presente exemplo, se λ é escolhido como sendo 0,1, então

$$\tau = 0{,}1t$$

$$\frac{dx}{dt} = 0{,}1\frac{dx}{d\tau}$$

$$\frac{d^2x}{dt^2} = 0{,}01\frac{d^2x}{d\tau^2}$$

A Eq. (6.54) agora se torna

$$\frac{d^2x}{d\tau^2} + 0{,}2\frac{dx}{d\tau} + 0{,}5x = \operatorname{sen} 10\tau$$

As condições iniciais são

$$x(0) = 2, \qquad \left.\frac{dx}{d\tau}\right|_{\tau=0} = 50$$

Note que em um problema particular, além da escala de tempo, pode também ser necessário escolher fatores de escala de magnitude de tal forma que a magnitude da função de excitação não seja nem muito grande nem muito pequena e que as saídas dos amplificadores tenham uma excursão cobrindo a maior parte de sua região de linearidade (±100 V).

Simulação de sistemas físicos em computador analógico. A simulação de sistemas dinâmicos é uma aplicação muito importante do computador analógico. Em particular, o computador analógico é muito útil para se determinar os efeitos de variações de parâmetros no desempenho do sistema.

Considere o sistema visto na Fig. 6.43. Queremos simular este sistema em um computador analógico e investigar o efeito de variações em a e b na resposta a degrau unitário.

A função de transferência de malha-fechada é

$$\frac{X(s)}{U(s)} = \frac{as + b}{s^3 + 3s^2 + (2 + a)s + b}$$

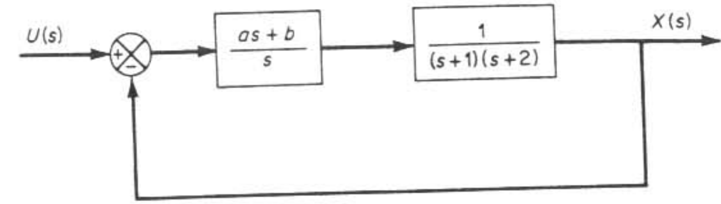


Fig. 6.43 Sistema de controle.

A equação diferencial correspondendo a esta função de transferência é

$$\dddot{x} + 3\ddot{x} + (2 + a)\dot{x} + bx = a\dot{u} + bu \tag{6.55}$$

Isolando no lado esquerdo a derivada de ordem mais alta da Eq. (6.55), obtemos

$$\dddot{x} = -3\ddot{x} - (2\dot{x} + a\dot{x} - a\dot{u}) - (bx - bu) \tag{6.56}$$

onde $x(0) = \dot{x}(0) = \ddot{x}(0) = 0$. Um diagrama de computação analógica, correspondendo à Eq. (6.56), é visto na Fig. 6.44. Este diagrama mostra apenas uma possibilidade. Se as gamas de variação dos valores de a e b são especificadas, podem-se determinar fatores de escala de magnitude adequados.

Os efeitos das constantes a e b na resposta a degrau unitário podem ser investigados variando-se o ajuste dos potenciômetros que envolvem a e b. (Se a, $b > 1$, poderemos necessitar da adição de amplificadores ao diagrama de computação.) As condições iniciais são ajustadas para zero.

Como outro exemplo da simulação de um sistema dinâmico, considere o sistema mecânico visto na Fig. 6.45(a). As equações

$$m_1\ddot{x}_1 + f_1\dot{x}_1 + k_1x_1 + k_2(x_1 - x_2) = 0$$

$$m_2\ddot{x}_2 + f_2\dot{x}_2 + k_3x_2 + k_2(x_2 - x_1) = 0$$

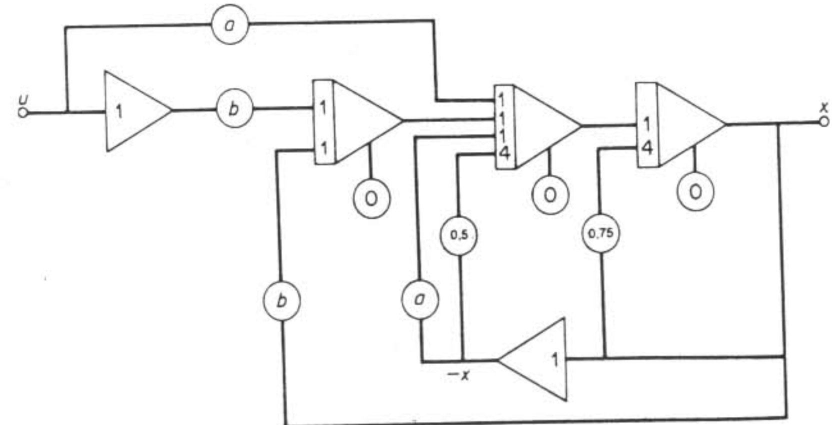
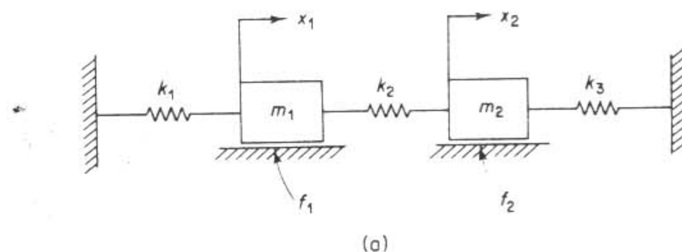


Fig. 6.44 Diagrama de computação analógica para resolver a Eq. (6.56).

Um diagrama de simulação em computador analógico para este sistema é visto na Fig. 6.45(b).

Comentários de conclusão. No decurso da análise e projeto de sistemas complicados, simulações em computadores analógicos têm um papel importante. Os efeitos de mudanças nos parâmetros do sistema sobre o desempenho deste podem ser facilmente determinados. A vantagem de simulações analógicas (ou digitais) reside no fato de que qualquer escala de tempo pode ser usada.

Em geral, a representação matemática precisa de um componente complicado é muito difícil. É provável que algumas das características importantes do componente não sejam incluídas na simulação (por desconhecimento). Isto pode causar



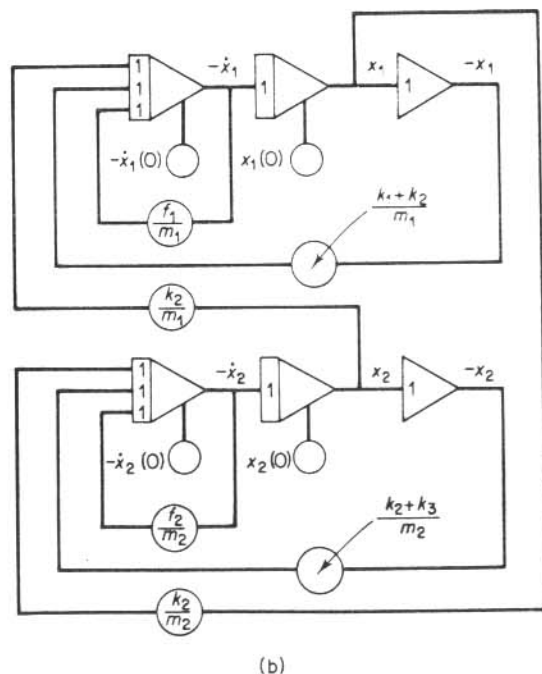


Fig. 6.45 (a) Sistema mecânico; (b) diagrama de simulação analógica para o sistema mecânico.

erros sérios na solução. Para evitar tais erros, o simulador pode incluir componentes do sistema real. Se tais componentes são incluídos, não se perde nenhuma característica importante do componente. A solução, entretanto, deve ser obtida em tempo real.

Apesar de ainda não termos discutido componentes não-lineares, o computador analógico é bastante versátil e conveniente de ser usado na simulação de sistemas de controle contendo tais componentes. Operações não-lineares como a multiplicação de duas variáveis podem ser feitas facilmente com o computador eletrônico analógico. Circuitos eletrônicos-padrão são disponíveis para simular não-linearidades comumente encontradas, como zona morta, saturação, histerese e atrito de Coulomb.

PROBLEMAS ILUSTRATIVOS E SOLUÇÕES

Problema A.6.1 Explique por que o controle proporcional de um processo que não possui uma propriedade de integração (o que significa que a função de transferência do processo não inclui o fator $1/s$) sofre de *offset* (desvio) em resposta a entradas em degrau.

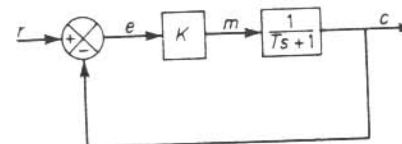


Fig. 6.46 Sistema de controle.

Solução. Considere, para exemplo, o sistema visto na Fig. 6.46. Em regime estacionário, se c fosse igual a uma constante não nula r, então $e = 0$ e $m = Ke = 0$, resultando em $c = 0$, o que contradiz a hipótese de que $c = r =$ constante não nula.

Um *offset* não nulo deve existir para a operação adequada de tal sistema de controle. Em outras palavras, em regime estacionário, se e fosse igual a $r/(1 + K)$, então $m = Kr/(1 + K)$ e $c = Kr/(1 + K)$, o que resulta no sinal de erro suposto $e = r/(1 + K)$. Portanto o *offset* é $r/(1 + K)$.

Problema A.6.2 Considere o sistema de controle de nível de líquido como visto na Fig. 6.47. O controlador é do tipo proporcional. A entrada de referência do controlador é fixa.

Desenhe um diagrama de blocos do sistema, supondo que as variações nas variáveis sejam pequenas. Investigue a resposta do nível do segundo tanque sujeito a uma perturbação em degrau u.

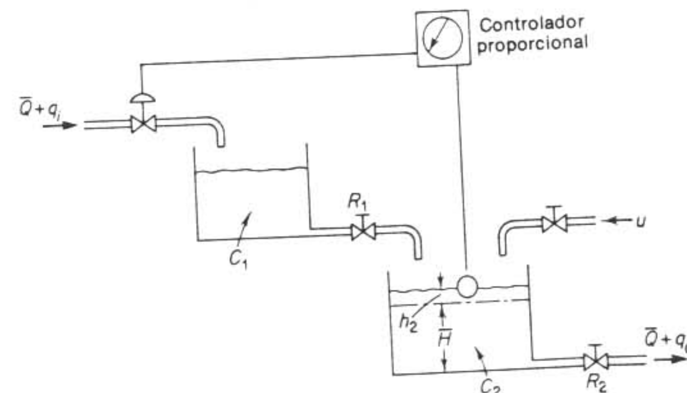


Fig. 6.47 Sistema de controle de nível de líquido.

Solução. A Fig. 6.48(a) é um diagrama de blocos deste sistema quando as variações nas variáveis são pequenas. Como a entrada de referência do controlador é fixa, $r = 0$. (Note que r é a variação na referência.)

Para investigar a resposta do nível do segundo tanque sujeito a uma perturbação em degrau u, achamos conveniente modificar o diagrama de blocos da Fig. 6.48(a) para aquele da Fig. 6.48(b).

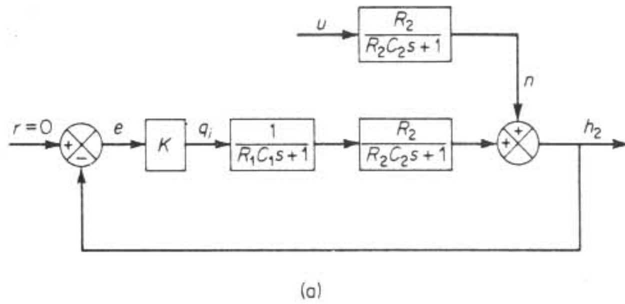
A função de transferência entre $H_2(s)$ e $U(s)$ pode ser obtida como

$$\frac{H_2(s)}{U(s)} = \frac{R_2(R_1C_1s + 1)}{(R_1C_1s + 1)(R_2C_2s + 1) + KR_2}$$

Desta equação, a resposta $H_2(s)$ para uma perturbação $U(s)$ pode ser determinada. O efeito do controlador é visto pela presença de K no denominador desta última equação.

Para uma perturbação em degrau de magnitude U_0, obtemos

$$h_2(\infty) = \frac{R_2}{1 + KR_2}U_0$$



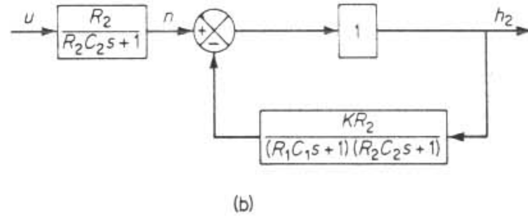


Fig. 6.48 (a) Diagrama de blocos do sistema visto na Fig. 6.47; (b) diagrama de blocos modificado.

ou

$$\text{Erro estacionário} = -\frac{R_2}{1 + KR_2}U_0$$

O sistema apresenta *offset* para uma perturbação em degrau. Na resposta a uma perturbação em degrau, tanto o coeficiente de amortecimento quanto a freqüência natural não-amortecida dependem do valor do ganho K. Este ganho deve ser ajustado de tal forma que a resposta transitória a perturbações apresente amortecimento razoável bem como velocidade razoável.

Problema A.6.3 Considere a resposta a degrau unitário do sistema de segunda-ordem

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

A amplitude da senóide amortecida exponencialmente varia seguindo uma série geométrica. No instante $t = t_p = \pi/\omega_d$, a amplitude é igual a $e^{-(\sigma/\omega_d)\pi}$. Após uma oscilação, ou em $t = t_p + 2\pi/\omega_d = 3\pi/\omega_d$, a amplitude é igual a $e^{-(\sigma/\omega_d)3\pi}$; após mais um ciclo de oscilação, a amplitude é $e^{-(\sigma/\omega_d)5\pi}$. O logaritmo desta relação de amplitude é o chamado decremento logarítmico. Determine o decremento logarítmico para este sistema de segunda-ordem.

Solução. A relação de amplitude por um período de oscilação amortecida é

$$\frac{c(t_1)}{c\left(t_1 + \dfrac{2\pi}{\omega_d}\right)} = \frac{e^{-(\sigma/\omega_d)\pi}}{e^{-(\sigma/\omega_d)3\pi}} = e^{2(\sigma/\omega_d)\pi} = e^{2\zeta\pi/\sqrt{1-\zeta^2}}$$

onde

$$t_1 = t_p + n\frac{2\pi}{\omega_d} \qquad (n = 0, 1, 2, \ldots)$$

Portanto, o decremento logarítmico δ é

$$\delta = \ln\left|\frac{c(t_1)}{c\left(t_1 + \dfrac{2\pi}{\omega_d}\right)}\right| = \frac{2\zeta\pi}{\sqrt{1-\zeta^2}}$$

Ele é uma função apenas do coeficiente de amortecimento ζ.

Problema A.6.4 Considere um sistema de controle com realimentação unitária cuja função de transferência de malha aberta é

$$G(s) = \frac{K}{s(Js + F)}$$

Discutir os efeitos da variação dos valores de K e F sobre o erro estacionário na resposta a rampa unitária. Esboce curvas de resposta a rampa unitária típica para um valor baixo, médio e alto para a constante K.

Solução. A função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{K}{Js^2 + Fs + K}$$

Para uma entrada rampa unitária, $R(s) = 1/s^2$. Portanto,

$$\frac{E(s)}{R(s)} = \frac{R(s) - C(s)}{R(s)} = \frac{Js^2 + Fs}{Js^2 + Fs + K}$$

ou

$$E(s) = \frac{Js^2 + Fs}{Js^2 + Fs + K}\frac{1}{s^2}$$

O erro estacionário é

$$e_{ss} = e(\infty) = \lim_{s \to 0} sE(s) = \frac{F}{K}$$

Vemos que podemos reduzir o erro estacionário e_{ss} aumentando o ganho K ou diminuindo o coeficiente de atrito viscoso F. Entretanto, aumentando o ganho ou diminuindo o coeficiente de atrito viscoso, obtém-se um decréscimo no coeficiente de amortecimento, resultando em uma resposta transitória mais oscilatória. Dobrando K, e_{ss} diminui para a metade do seu valor original, enquanto que ζ é diminuído para 0,707 do seu valor original visto que ζ é inversamente proporcional à raiz quadrada de K. Por outro lado, diminuindo F para metade do seu valor original, tanto e_{ss} como ζ diminuem para as metades dos seus valores originais, respectivamente. Portanto, é recomendável aumentar o valor de K ao invés de diminuir o valor de F. Após o transitório ter desaparecido e ter-se alcançado um estado estacionário, a velocidade de saída se torna igual à velocidade de entrada. Entretanto, há um erro posicional estacionário entre a entrada e a saída. Exemplos da resposta a rampa unitária do sistema para três valores distintos de K são ilustrados na Fig. 6.49.

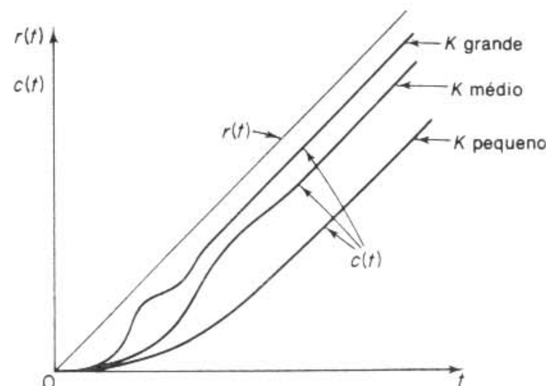


Fig. 6.49 Curvas de resposta a rampa unitária do sistema considerado no Problema A.6.4.

Problema A.6.5 Obtenha a resposta a rampa unitária do sistema dado por

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

Obtenha também o erro estacionário.

Solução. Para uma entrada rampa unitária, $R(s) = 1/s^2$. A saída $C(s)$ pode ser escrita

$$C(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}\frac{1}{s^2}$$

Para o caso subamortecido ($0 \leq \zeta < 1$), $c(t)$, a transformada inversa de Laplace de $C(s)$, pode ser obtida como

$$c(t) = t - \frac{2\zeta}{\omega_n} + e^{-\zeta\omega_n t}\left(\frac{2\zeta}{\omega_n}\cos\omega_d t + \frac{2\zeta^2 - 1}{\omega_n\sqrt{1-\zeta^2}}\operatorname{sen}\omega_d t\right)$$

$$= t - \frac{2\zeta}{\omega_n} + \frac{e^{-\zeta\omega_n t}}{\omega_n\sqrt{1-\zeta^2}}\operatorname{sen}\left(\omega_d t + \tan^{-1}\frac{2\zeta\sqrt{1-\zeta^2}}{2\zeta^2 - 1}\right) \qquad (t \geq 0)$$

onde

$$\omega_d = \omega_n\sqrt{1-\zeta^2}$$

$$\tan^{-1}\frac{2\zeta\sqrt{1-\zeta^2}}{2\zeta^2 - 1} = 2\tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}$$

Para o caso de amortecimento crítico ($\zeta = 1$), $c(t)$ é

$$c(t) = t - \frac{2\zeta}{\omega_n} + \frac{2}{\omega_n}e^{-\omega_n t}\left(1 + \frac{\omega_n t}{2}\right) \qquad (t \geq 0)$$

Para o caso sobreamortecido ($\zeta > 1$), $c(t)$ é

$$c(t) = t - \frac{2\zeta}{\omega_n} - \frac{2\zeta^2 - 1 - 2\zeta\sqrt{\zeta^2 - 1}}{2\omega_n\sqrt{\zeta^2 - 1}}e^{-(\zeta + \sqrt{\zeta^2 - 1})\omega_n t}$$

$$+ \frac{2\zeta^2 - 1 + 2\zeta\sqrt{\zeta^2 - 1}}{2\omega_n\sqrt{\zeta^2 - 1}}e^{-(\zeta - \sqrt{\zeta^2 - 1})\omega_n t} \qquad (t \geq 0)$$

(A resposta a rampa unitária deste sistema pode também ser obtida integrando-se a resposta a degrau unitário com relação ao tempo e determinando as constantes de integração, usando-se a condição de que as condições iniciais na saída são nulas.) O sinal de erro $e(t)$ neste caso é

$$e(t) = r(t) - c(t) = t - c(t)$$

O erro estacionário, e_{ss}, para $\zeta > 0$ é então

$$e_{ss} = e(\infty) = \frac{2\zeta}{\omega_n}$$

Problema A.6.6 Considere o sistema visto na Fig. 6.50(a). O erro estacionário para uma entrada rampa unitária é $e_{ss} = 2\zeta/\omega_n$. Mostre que o erro estacionário para seguir uma entrada em rampa pode ser eliminado se a entrada é introduzida no sistema através de um elemento tipo proporcional-mais-derivada, como visto na Fig. 6.50(b), com o valor de k adequadamente escolhido.

Solução. A função de transferência de malha-fechada do sistema visto na Fig. 6.50(b) é

$$\frac{C(s)}{R(s)} = \frac{(1 + ks)\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

Então

$$R(s) - C(s) = \left(\frac{s^2 + 2\zeta\omega_n s - \omega_n^2 ks}{s^2 + 2\zeta\omega_n s + \omega_n^2}\right)R(s)$$

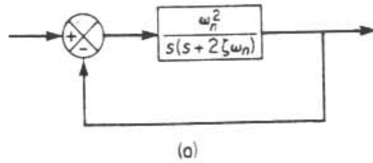
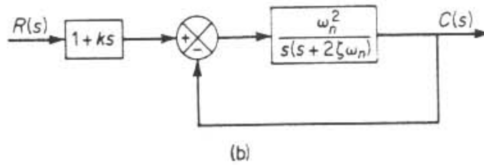


Fig. 6.50 Sistemas de controle.



Se a entrada for uma rampa unitária, então o erro estacionário é

$$\begin{aligned} e(\infty) &= r(\infty) - c(\infty) \\ &= \lim_{s\to 0} s\left(\frac{s^2 + 2\zeta\omega_n s - \omega_n^2 ks}{s^2 + 2\zeta\omega_n s + \omega_n^2}\right)\frac{1}{s^2} \\ &= \frac{2\zeta\omega_n - \omega_n^2 k}{\omega_n^2} \end{aligned}$$

Portanto, se k é escolhido como sendo

$$k = \frac{2\zeta}{\omega_n}$$

então o erro estacionário ao se seguir uma entrada em rampa pode ser feito igual a zero. Observe que se houver quaisquer variações nos valores de ζ e/ou ω_n devido a variações ambientais ou envelhecimento, então um erro estacionário não nulo poderá resultar para uma resposta a rampa.

Problema A.6.7 Considere o servomecanismo visto na Fig. 6.51. Determine os valores de K e k de tal forma que o sobre-sinal máximo na resposta a degrau unitário seja de 25% e o instante do pico seja 2 s.

Solução. O sobre-sinal máximo M_p é

$$M_p = e^{-(\zeta/\sqrt{1-\zeta^2})\pi}$$

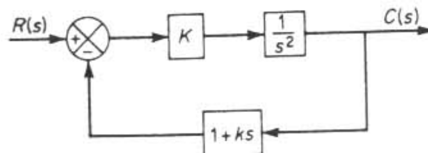


Fig. 6.51 Diagrama de blocos de um servomecanismo.

que é especificado como sendo 25%. Portanto,

$$e^{-(\zeta/\sqrt{1-\zeta^2})\pi} = 0{,}25$$

donde temos que

$$\frac{\zeta}{\sqrt{1-\zeta^2}}\pi = 1{,}39$$

ou

$$\zeta = 0{,}4$$

O instante do pico t_p é especificado como sendo 2 s. Portanto,

$$t_p = \frac{\pi}{\omega_d} = 2$$

ou

$$\omega_d = 1{,}57$$

Então, a freqüência natural não amortecida ω_n é

$$\omega_n = \frac{\omega_d}{\sqrt{1-\zeta^2}} = \frac{1{,}57}{\sqrt{1-0{,}4^2}} = 1{,}71$$

Do diagrama de blocos, obtemos

$$\frac{C(s)}{R(s)} = \frac{K}{s^2 + Kks + K}$$

Observando que

$$\omega_n = \sqrt{K}, \qquad \zeta = \frac{Kk}{2\omega_n}$$

obtemos

$$\begin{aligned} K &= \omega_n^2 = 1{,}71^2 = 2{,}93 \\ k &= \frac{2\zeta\omega_n}{K} = \frac{2 \times 0{,}4 \times 1{,}71}{2{,}93} = 0{,}47 \end{aligned}$$

Problema A.6.8 A Fig. 6.52(a) mostra um sistema mecânico vibratório. Quando uma força (entrada em degrau) de 2 lb é aplicada ao sistema, a massa oscila, como visto na Fig. 6.52(b). Determine m, f e k do sistema a partir desta curva de resposta.

Solução. A função de transferência deste sistema é

$$\frac{X(s)}{P(s)} = \frac{1}{ms^2 + fs + k}$$

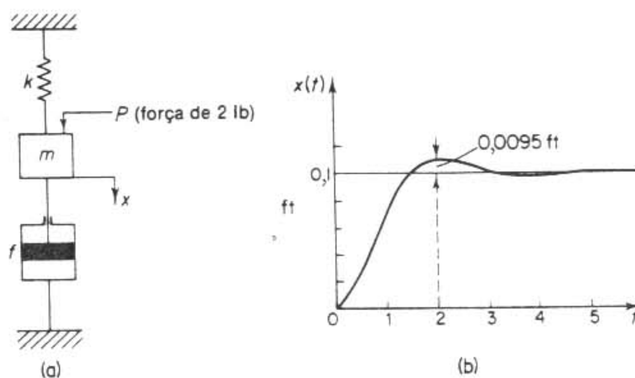


Fig. 6.52 (a) Sistema mecânico vibratório; (b) curva de resposta a degrau.

Como

$$P(s) = \frac{2}{s}$$

obtemos

$$X(s) = \frac{2}{s(ms^2 + fs + k)}$$

Segue-se então que o valor estacionário de x é

$$x(\infty) = \lim_{s \to 0} sX(s) = \frac{2}{k} = 0{,}1 \text{ ft}$$

Portanto

$$k = 20 \text{ lb/ft}$$

Note que M_p = 9,5% corresponde a ζ = 0,6. O instante do pico t_p é dado por

$$t_p = \frac{\pi}{\omega_d} = \frac{\pi}{\omega_n\sqrt{1-\zeta^2}} = \frac{\pi}{0{,}8\omega_n}$$

A curva experimental mostra que t_p = 2 s. Portanto,

$$\omega_n = \frac{3{,}14}{2 \times 0{,}8} = 1{,}96 \text{ rad/s}$$

Como $\omega_n^2 = k/m = 20/m$, obtemos

$$m = \frac{20}{\omega_n^2} = \frac{20}{1{,}96^2} = 5{,}2 \text{ slugs}$$
$$= 166 \text{ lb}$$

Então f é determinado a partir de

$$2\zeta\omega_n = \frac{f}{m}$$

ou

$$f = 2\zeta\omega_n m = 2 \times 0{,}6 \times 1{,}96 \times 5{,}2 = 12{,}2 \text{ lb/ft/s}$$

Problema A.6.9 Considere um sistema cujos pólos de malha-fechada e zeros de malha-fechada estão localizados no plano s em uma linha paralela ao eixo $j\omega$, como visto na Fig. 6.53. Mostre que a resposta impulsiva de um sistema como este é uma função co-senoidal amortecida.

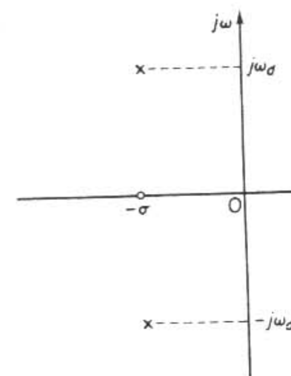


Fig. 6.53 Configuração de pólos e zeros de malha-fechada.

Solução. A função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{K(s+\sigma)}{(s+\sigma+j\omega_d)(s+\sigma-j\omega_d)}$$

Para uma entrada impulso unitário, $R(s)$ = 1 e

$$C(s) = \frac{K(s+\sigma)}{(s+\sigma)^2 + \omega_d^2}$$

A transformada inversa de Laplace de $C(s)$ é

$$c(t) = Ke^{-\sigma t}\cos\omega_d t \qquad (t \geq 0)$$

que é uma função co-senoidal amortecida.

Problema A.6.10 Considere o sistema visto na Fig. 6.54. Este sistema é excitado por doi sinais, um a entrada de referência e o outro a perturbação externa. Mostre que a equaçã característica deste sistema é a mesma, independentemente de qual sinal é escolhido com entrada.

Solução. A função de transferência que relaciona a entrada de referência e a saída correspor

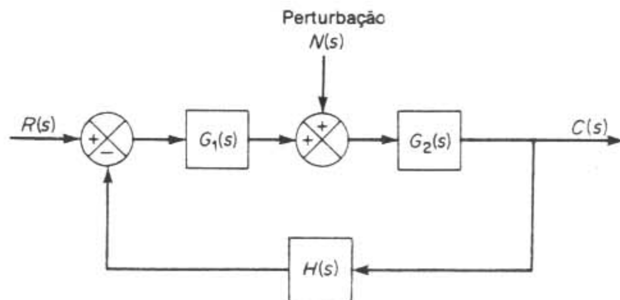


Fig. 6.54 Sistema de controle.

dente, sem considerar a perturbação externa, é

$$\frac{C(s)}{R(s)} = \frac{G_1 G_2}{1 + G_1 G_2 H} \tag{6.57}$$

A função de transferência que relaciona a perturbação externa e a correspondente saída na ausência da entrada de referência é

$$\frac{C(s)}{N(s)} = \frac{G_2}{1 + G_1 G_2 H} \tag{6.58}$$

Note que os denominadores das Eqs. (6.57) e (6.58) são os mesmos. A equação característica é

$$1 + G_1 G_2 H = 0$$

Esta contém a informação necessária para determinar características básicas da resposta do sistema. (Lembre-se que para um dado sistema há apenas uma equação característica. Isto significa que a equação característica de uma dada função de transferência é a mesma independentemente de qual sinal é escolhido como entrada.)

Problema A.6.11 Os zeros de malha-aberta e os zeros de malha-fechada são idênticos para um sistema a malha-fechada?

Solução. Não. Considere um sistema a malha-fechada cuja função de transferência do elo direto é $G(s) = p(s)/q(s)$ e a função de transferência do elo de realimentação é $H(s) = n(s)/d(s)$, onde $p(s)$, $q(s)$, $n(s)$ e $d(s)$ são polinômios em s. Então,

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} = \frac{p(s)d(s)}{q(s)d(s) + p(s)n(s)}$$

Vemos que os zeros da função de transferência de malha-fechada são aqueles valores de s que fazem $p(s)d(s) = 0$ e que os zeros da função de transferência de malha-aberta são aqueles valores de s que fazem $p(s)n(s) = 0$. (Portanto, alguns dos zeros da função de transferência de malha-fechada são os mesmos dos da função de transferência de malha-aberta.)

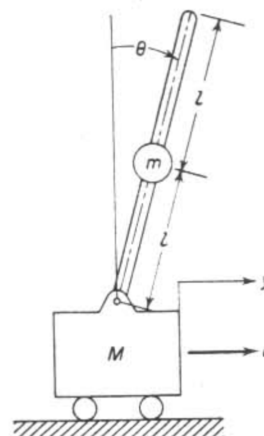


Fig. 6.55 Pêndulo invertido montado em um carrinho motorizado.

Problema A.6.12* Considere o pêndulo invertido montado em um carrinho motorizado visto na Fig. 6.55. Este é um modelo do controle de atitude de um foguete por ocasião do lançamento. O objetivo do problema do controle de atitude é manter o foguete na posição vertical. No problema presente, queremos manter o pêndulo na posição vertical. O foguete (ou o pêndulo invertido neste problema) é instável e pode cair a qualquer instante e em qualquer direção.

Neste problema, consideramos apenas um problema bidimensional de tal forma que o pêndulo visto na Fig. 6.55 se move apenas no plano da folha do livro. Para manter o pêndulo invertido na vertical, suponha que medimos θ e $\dot{\theta}$ continuamente e que usamos um controlador do tipo proporcional-mais-derivada para produzir a força de controle u, ou

$$u = M(a\theta + b\dot{\theta})$$

Determine as condições para a e b de tal forma que o sistema seja estável. Suponha que não há atrito no eixo ligando o pêndulo ao carrinho nem nas rodas do carrinho. Suponha também que θ e $\dot{\theta}$ são pequenos.

Solução. Usando os símbolos definidos na Fig. 6.55, podemos obter as equações diferenciais descrevendo a dinâmica do sistema como segue: Observando que θ e $\dot{\theta}$ são pequenos, de tal forma que sen $\theta \doteqdot \theta$, cos $\theta \doteqdot 1$, temos

$$(J + ml^2)\ddot{\theta} + ml\ddot{y} - mgl\theta = 0 \tag{6.59}$$

$$ml\ddot{\theta} + (M + m)\ddot{y} = u \tag{6.60}$$

$$M(a\theta + b\dot{\theta}) = u \tag{6.61}$$

onde J é o momento de inércia do pêndulo invertido relativo ao centro de massa. Como J pode ser dado por

$$J = \frac{ml^2}{3}$$

*O problema de controlar sistemas instáveis como o pêndulo invertido montado em um carrinho motorizado é visto com detalhes na Referência H-6.

a Eq. (6.59) pode ser escrita

$$\frac{4}{3}ml^2\ddot{\theta} + ml\ddot{y} - mgl\theta = 0$$

do qual

$$\ddot{y} = -\frac{4}{3}l\ddot{\theta} + g\theta \qquad (6.62)$$

Substituindo as Eqs. (6.62) e (6.61) na Eq. (6.60), temos

$$ml\ddot{\theta} + (M + m)\left(-\frac{4}{3}l\ddot{\theta} + g\theta\right) = M(a\theta + b\dot{\theta}) \qquad (6.63)$$

Vamos definir

$$\frac{m}{M} = \mu$$

Então, a Eq. (6.63) se torna

$$\ddot{\theta} + \frac{3b}{(\mu + 4)l}\dot{\theta} + \frac{3[a - (1 + \mu)g]}{(\mu + 4)l}\theta = 0$$

Para haver estabilidade, os coeficientes de θ e $\dot{\theta}$ devem ser positivos. Portanto, as condições de estabilidade são

$$b > 0, \qquad a > (1 + \mu)g$$

Se os valores de a e b são escolhidos satisfazendo estas condições, qualquer pequeno desvio na inclinação pode ser recuperado sem causar a queda do pêndulo.

Problema A.6.13 Determine a resposta a degrau unitário de um sistema com realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{5(s + 20)}{s(s + 4{,}59)(s^2 + 3{,}41s + 16{,}35)}$$

Solução. A função de transferência de malha-fechada é

$$\begin{aligned}\frac{C(s)}{R(s)} &= \frac{5(s + 20)}{s(s + 4{,}59)(s^2 + 3{,}41s + 16{,}35) + 5(s + 20)} \\ &= \frac{5s + 100}{s^4 + 8s^3 + 32s^2 + 80s + 100} \\ &= \frac{5(s + 20)}{(s^2 + 2s + 10)(s^2 + 6s + 10)}\end{aligned}$$

A resposta a degrau unitário deste sistema é então

$$\begin{aligned}C(s) &= \frac{5(s + 20)}{s(s^2 + 2s + 10)(s^2 + 6s + 10)} \\ &= \frac{1}{s} + \frac{\frac{3}{8}(s + 1) - \frac{17}{8}}{(s + 1)^2 + 3^2} + \frac{-\frac{11}{8}(s + 3) - \frac{13}{8}}{(s + 3)^2 + 1^2}\end{aligned}$$

A resposta temporal $c(t)$ pode ser achada tomando-se a transformada inversa de Laplace de $C(s)$ como segue:

$$c(t) = 1 + \tfrac{3}{8}e^{-t}\cos 3t - \tfrac{17}{24}e^{-t}\,\mathrm{sen}\,3t - \tfrac{11}{8}e^{-3t}\cos t - \tfrac{13}{8}e^{-3t}\,\mathrm{sen}\,t \qquad (t \geq 0)$$

PROBLEMAS

Problema B.6.1 Um termômetro requer 1 min para indicar 98% da resposta a uma entrada em degrau. Supondo que o termômetro é um sistema de primeira-ordem, determine a constante de tempo.

Se o termômetro é colocado em um banho, cuja temperatura varia linearmente a uma taxa de 10°/min, quanto erro haverá na leitura do termômetro?

Problema B.6.2 Qual seria o erro se o circuito visto na Fig. 6.56 fosse usado como um integrador?

Problema B.6.3 Determine a resposta a degrau unitário de um sistema com realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{4}{s(s + 5)}$$

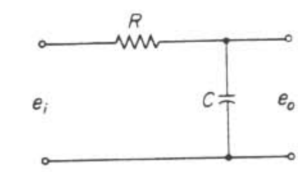


Fig. 6.56 Sistema elétrico.

Problema B.6.4 Considere a resposta a degrau unitário de um sistema de controle com realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{1}{s(s + 1)}$$

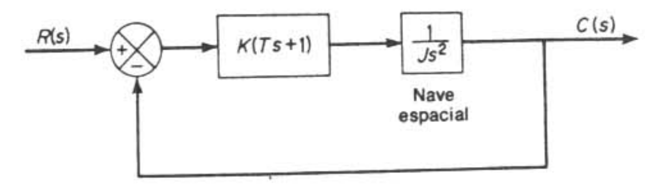


Fig. 6.57 Sistema de controle de atitude de uma nave espacial.

Determine o tempo de subida, o instante do pico, o sobre-sinal máximo, e o tempo de acomodação.

Problema B.6.5 Considere o sistema a malha-fechada dado por

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

Determine os valores de ζ e ω_n de tal forma que o sistema responda a uma entrada em degrau com aproximadamente 5% de sobre-sinal e com um tempo de acomodação de 2 s. (Usar o critério de 2%.)

Problema B.6.6 A Fig. 6.57 é um diagrama de blocos de um sistema de controle de atitude de uma nave espacial. Supondo a constante de tempo T do controlador como 3 s e a relação do torque para inércia K/J como 2/9 rad²/s², determine o coeficiente de amortecimento do sistema.

Problema B.6.7 Considere um sistema de controle com realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{0{,}4s + 1}{s(s + 0{,}6)}$$

Determine a resposta a entrada degrau unitário. Qual é o tempo de subida deste sistema? Qual é o sobre-sinal máximo?

Problema B.6.8 Determine a resposta a impulso unitário e a reposta a degrau unitário de um sistema com realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{2s + 1}{s^2}$$

Problema B.6.9 Considere o sistema visto na Fig. 6.58. Mostre que a função de transferência $Y(s)/X(s)$ tem um zero no semiplano direito do plano s. Em seguida obtenha $y(t)$ quando $x(t)$ é um degrau unitário. Faça um gráfico de $y(t)$ em função de t.

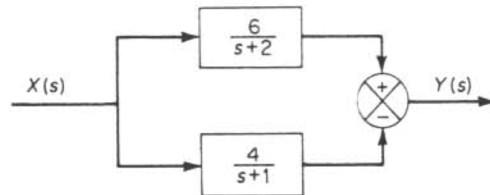


Fig. 6.58 Sistema com um zero no semiplano direito do plano s (Sistema não de mínima fase.)

Problema B.6.10 Aplique o critério de estabilidade de Routh para a seguinte equação:

$$s^4 + Ks^3 + s^2 + s + 1 = 0$$

e determine a gama de valores de K para haver estabilidade.

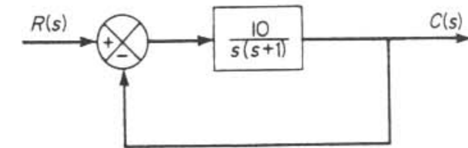
Problema B.6.11 Determine a gama de valores de K para haver estabilidade de um sistema de controle a realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{K}{s(s + 1)(s + 2)}$$

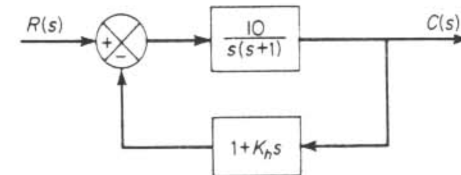
Problema B.6.12 Os dois sistemas abaixo são estáveis?

1. $G(s) = \dfrac{10(s + 1)}{s(s - 1)(s + 5)}$, $\quad H(s) = 1$
2. $G(s) = \dfrac{10}{s(s - 1)(2s + 3)}$, $\quad H(s) = 1$

Problema B.6.13 Considere o sistema visto na Fig. 6.59(a). O coeficiente de amortecimento deste sistema é 0,137 e a freqüência natural não amortecida é 3,16 rad/s. Para melhorar a estabilidade relativa, usamos uma realimentação por tacômetro. A Fig. 6.59(b) mostra o sistema com realimentação com tacômetro.



(a)



(b)

Fig. 6.59 (a) Sistema de controle; (b) sistema de controle com realimentação por tacômetro.

Determine o valor de K_h de tal forma que o coeficiente de amortecimento do sistema seja 0,5. Desenhe curvas de resposta a degrau unitário tanto para o sistema original como para o sistema com realimentação a tacômetro. Desenhe também as curvas de erro em função do tempo para a resposta a rampa unitária para ambos os sistemas.

Problema B.6.14 Faça um diagrama de computação analógica para o sistema de mola-massa-viscosidade descrito por

$$\ddot{x} + 2\dot{x} + 3x = u(t)$$

onde

$$x(0) = 2 \text{ ft}, \qquad \dot{x}(0) = 5 \text{ ft/s}$$

A entrada $u(t)$ é uma função degrau e é dada por

$$u(t) = 10 \text{ lb}$$

Problema B.6.15 Suponha que a característica dinâmica de um processo é dada por

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \qquad (0 < \zeta < 1)$$

A resposta deste sistema a uma entrada em degrau $a\,1(t)$ bem como para uma entrada em degrau atrasado $b\,1(t - t_1)$ é vista nas Figs. 6.60(a) e 6.60(b), respectivamente. Podemos ver a partir destas curvas de resposta que, se escolhermos os valores de a, b e t_1 de forma adequada, podemos então obter a resposta com tempo de acomodação finito sem oscilação como visto na Fig. 6.60(c). O método de controle de se aplicar uma entrada em degrau e uma entrada em degrau atrasado para obter uma resposta com tempo de acomodação finito e sem oscilação é chamado *controle posicast*. (Este nome foi dado por O.J.M. Smith.)

Se a característica dinâmica de um processo pode ser representada por uma função de transferência de segunda-ordem subamortecida, então este método pode ser aplicado convenientemente para ajustar a entrada de referência do controlador para iniciar o ciclo do processo.

Suponha que desejamos iniciar o processo e levar a saída $c(t)$ para h através de controle *posicast*. Determine os valores de a, b e t_1 vistos na Fig. 6.60(c) em termos de h, ζ e ω_n.

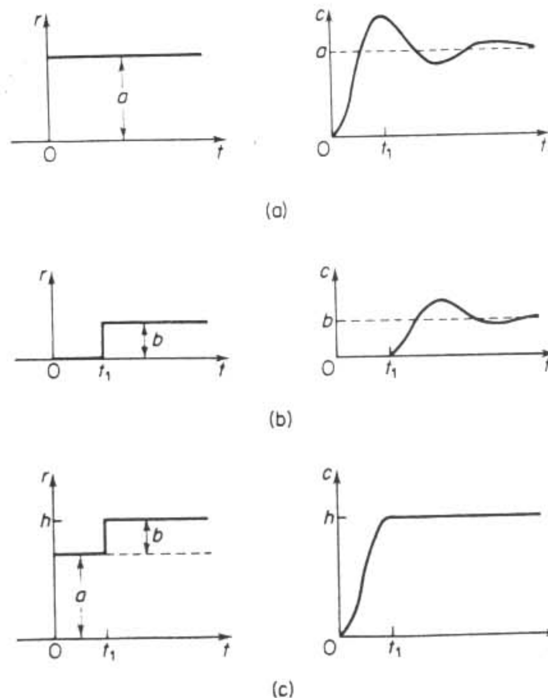


Fig. 6.60 (a) Entrada em degrau e curva de resposta; (b) entrada em degrau atrasado e curva de resposta; (c) curvas de entrada e resposta mostrando a característica de controle *posicast*.

7

Análise de Erros e Introdução à Otimização de Sistemas

7.1 COEFICIENTES DE ERRO ESTÁTICO

As características de resposta transitória discutidas no Cap. 6 são facetas importantes do desempenho de sistemas de controle. Uma outra característica importante se relaciona com o erro do sistema. Os erros em um sistema de controle podem ser atribuídos a muitos fatores. Variações na entrada de referência causarão erros inevitáveis durante períodos transitórios e podem também ocasionar erros estacionários. Imperfeições nos componentes do sistema, como por exemplo, atrito estático, folga *(backlash)*, e deriva de amplificadores, bem como envelhecimento ou deterioração, causarão erros em regime estacionário.

Em capítulos anteriores, foi dito que o erro estacionário é uma medida da precisão de um sistema de controle. O desempenho em regime estacionário de um sistema de controle estável é geralmente julgado pelo erro estacionário devido a entradas em degrau, rampa, ou aceleração. Nesta seção, investigaremos tal desempenho em regime estacionário. Mais precisamente, pesquisaremos um tipo de erro estacionário que é causado pela incapacidade de um sistema de seguir certos tipos de entradas.

Qualquer sistema de controle físico sofre inerentemente de erro estacionário em resposta a certos tipos de entradas. Um sistema pode não apresentar erro estacionário para entradas em degrau, mas o mesmo sistema pode apresentar erro estacionário não nulo para uma entrada em rampa. (A única maneira de se eliminar este erro é modificando a estrutura do sistema.) O fato de um dado sistema apresentar erro estacionário para um dado tipo de entrada depende do tipo de função de transferência de malha-aberta do sistema, o que será discutido a seguir.

Classificações de sistemas de controle. Sistemas de controle podem ser classificados de acordo com a sua habilidade para seguir entradas em degrau, entradas em

rampa, entradas parabólicas etc. Este é um esquema razoável de classificação porque as entradas reais podem ser freqüentemente consideradas como combinações de tais entradas. Os valores dos erros estacionários devidos a estas entradas individuais são indicativos da qualidade do sistema.

Considere a seguinte função de transferência de malha-aberta $G(s)H(s)$:

$$G(s)H(s) = \frac{K(T_a s + 1)(T_b s + 1) \cdots (T_m s + 1)}{s^N(T_1 s + 1)(T_2 s + 1) \cdots (T_p s + 1)}$$

Ela envolve o termo s^N no denominador, representando um pólo de multiplicidade N na origem. O presente esquema de classificação é baseado no número de integrações indicadas pela função de transferência de malha-aberta. Um sistema é chamado de tipo 0, tipo 1, tipo 2, ... se $N = 0, N = 1, N = 2, \ldots$, respectivamente. Note que esta classificação é diferente daquela de ordem de um sistema. Aumentando-se o número do tipo, melhora a precisão mas piora o problema da estabilidade. Um compromisso entre precisão em regime estacionário e estabilidade relativa sempre é necessário. Na prática, raramente se tem um sistema de tipo 3 ou maior porque geralmente é difícil projetar sistemas estáveis com mais do que duas integrações no ramo direto.

Veremos mais tarde que, se $G(s)H(s)$ é escrita de tal forma que cada termo no numerador e denominador, exceto o termo s^N, tende à unidade quando s tende a zero, então o ganho de malha-aberta K está diretamente relacionada com o erro estacionário.

Erros estacionários. Considere o sistema visto na Fig. 7.1. A função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)}$$

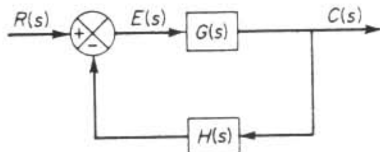


Fig. 7.1 Sistema de controle.

A função de transferência entre o sinal de erro atuante $e(t)$ e o sinal de entrada $r(t)$ é

$$\frac{E(s)}{R(s)} = 1 - \frac{C(s)H(s)}{R(s)} = \frac{1}{1 + G(s)H(s)}$$

onde o erro atuante $e(t)$ é a diferença entre o sinal de entrada e o sinal de realimentação.

O teorema do valor final provê uma maneira conveniente de se determinar o desempenho estacionário de um sistema estável. Como $E(s)$ é

$$E(s) = \frac{1}{1 + G(s)H(s)} R(s)$$

o erro atuante estacionário é

$$e_{ss} = \lim_{t\to\infty} e(t) = \lim_{s\to 0} \frac{sR(s)}{1 + G(s)H(s)}$$

Os coeficientes de erro estático definidos a seguir são figuras de mérito de sistemas de controle. Quanto maiores os coeficientes, menor o erro estacionário. Em um dado sistema, a saída pode ser posição, velocidade, pressão, temperatura etc. A forma física da saída é, entretanto, imaterial nesta análise. Portanto, daqui em diante chamaremos a saída de "posição", a taxa de variação da saída de "velocidade" etc. Isto significa que em um sistema de controle de temperatura "posição" representa a temperatura de saída, "velocidade" representa a taxa de variação da temperatura de saída etc.

Coeficiente de erro de posição estático K_p. O erro atuante estacionário do sistema para uma entrada degrau unitário é

$$e_{ss} = \lim_{s\to 0} \frac{s}{1 + G(s)H(s)} \frac{1}{s}$$
$$= \frac{1}{1 + G(0)H(0)}$$

O coeficiente de erro de posição estático K_p é definido por

$$K_p = \lim_{s\to 0} G(s)H(s) = G(0)H(0)$$

Portanto, o erro atuante estacionário é dado em termos do coeficiente de erro de posição estático K_p pela expressão

$$e_{ss} = \frac{1}{1 + K_p}$$

Para um sistema do tipo 0,

$$K_p = \lim_{s\to 0} \frac{K(T_a s + 1)(T_b s + 1) \cdots}{(T_1 s + 1)(T_2 s + 1) \cdots} = K$$

Para um sistema do tipo 1 ou maior,

$$K_p = \lim_{s\to 0} \frac{K(T_a s + 1)(T_b s + 1) \cdots}{s^N(T_1 s + 1)(T_2 s + 1) \cdots} = \infty \qquad (N \geq 1)$$

Portanto, para um sistema do tipo 0, o coeficiente de erro de posição estático K_p é finito, enquanto que para um sistema do tipo 1 ou maior K_p é infinito.

Para uma entrada degrau unitário, o erro atuante estacionário e_{ss} pode ser resumido como segue:

$$e_{ss} = \frac{1}{1 + K} \qquad \text{para sistemas do tipo 0}$$

$$e_{ss} = 0 \qquad \text{para sistemas do tipo 1 ou maior}$$

Da análise feita, vemos que a resposta de um sistema de controle com realimentação unitária para uma entrada em degrau envolve um erro estacionário se não há integração no ramo direto. (Se erros pequenos podem ser tolerados para entradas em degrau, então um sistema do tipo 0 pode ser admissível, contanto que o ganho K seja suficientemente grande. Entretanto, se o ganho K é grande demais, fica difícil se obter uma estabilidade relativa adequada.) Se desejarmos um erro estacionário nulo para uma entrada em degrau, então o tipo do sistema deve ser 1 ou maior.

Coeficiente de erro de velocidade estático K_v. O erro atuante estacionário do sistema com uma entrada rampa unitária (entrada velocidade unitária) é dado por

$$e_{ss} = \lim_{s\to 0} \frac{s}{1 + G(s)H(s)} \frac{1}{s^2}$$
$$= \lim_{s\to 0} \frac{1}{sG(s)H(s)}$$

O coeficiente de erro de velocidade estático K_v é definido por

$$K_v = \lim_{s\to 0} sG(s)H(s)$$

Portanto, o erro atuante estacionário é dado em termos do coeficiente de erro de velocidade estático K_v pela expressão

$$e_{ss} = \frac{1}{K_v}$$

O termo *erro de velocidade* é usado aqui para significar o erro para uma entrada em rampa. A dimensão do erro de velocidade é a mesma que a do erro do sistema, isto é, o erro de velocidade não é um erro na velocidade, mas um erro na posição devido a uma entrada em rampa.

Para um sistema do tipo 0,

$$K_v = \lim_{s\to 0} \frac{sK(T_a s + 1)(T_b s + 1)\cdots}{(T_1 s + 1)(T_2 s + 1)\cdots} = 0$$

Para um sistema do tipo 1,

$$K_v = \lim_{s\to 0} \frac{sK(T_a s + 1)(T_b s + 1)\cdots}{s(T_1 s + 1)(T_2 s + 1)\cdots} = K$$

Para um sistema do tipo 2 ou maior,

$$K_v = \lim_{s\to 0} \frac{sK(T_a s + 1)(T_b s + 1)\cdots}{s^N(T_1 s + 1)(T_2 s + 1)\cdots} = \infty \qquad (N \geq 2)$$

O erro atuante estacionário e_{ss} para a entrada rampa unitária pode ser resumido como segue:

$$e_{ss} = \frac{1}{K_v} = \infty \qquad \text{para sistemas do tipo 0}$$
$$e_{ss} = \frac{1}{K_v} = \frac{1}{K} \qquad \text{para sistemas do tipo 1}$$
$$e_{ss} = \frac{1}{K_v} = 0 \qquad \text{para sistemas do tipo 2 ou maior}$$

A análise feita mostra que um sistema do tipo 0 é incapaz de seguir uma entrada em rampa no regime estacionário. O sistema do tipo 1 com realimentação unitária pode seguir a entrada em rampa com um erro finito. Em operação estacionária, a velocidade de saída é exatamente igual à velocidade de entrada, mas há um erro na posição. Este erro é proporcional à velocidade da entrada e inversamente proporcional ao ganho K. A Fig. 7.2 mostra um exemplo da resposta de um sistema do tipo 1 com realimentação unitária para entrada em rampa. O sistema de ordem 2 ou maior pode seguir uma entrada em rampa com erro atuante nulo em regime estacionário.

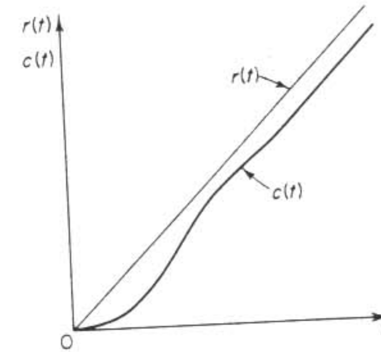


Fig. 7.2 Resposta de um sistema com realimentação unitária, do tipo 1, para uma entrada em rampa.

Coeficiente de erro de aceleração estático K_a. O erro atuante estacionário do sistema com uma entrada parábola unitária (entrada de aceleração) que é definida por

$$r(t) = \frac{t^2}{2} \qquad \text{para } t \geq 0$$
$$= 0 \qquad \text{para } t < 0$$

é dado por

$$e_{ss} = \lim_{s\to 0} \frac{s}{1 + G(s)H(s)} \frac{1}{s^3} = \frac{1}{\lim_{s\to 0} s^2 G(s)H(s)}$$

O coeficiente de erro de aceleração estático K_a é definido pela equação

$$K_a = \lim_{s \to 0} s^2 G(s)H(s)$$

O erro atuante estacionário é então

$$e_{ss} = \frac{1}{K_a}$$

Note que o erro de aceleração, o erro estacionário devido a uma entrada em parábola, é um erro na posição.

Os valores de K_a são obtidos como segue:

Para um sistema do tipo 0,

$$K_a = \lim_{s \to 0} \frac{s^2 K(T_a s + 1)(T_b s + 1) \cdots}{(T_1 s + 1)(T_2 s + 1) \cdots} = 0$$

Para um sistema do tipo 1,

$$K_a = \lim_{s \to 0} \frac{s^2 K(T_a s + 1)(T_b s + 1) \cdots}{s(T_1 s + 1)(T_2 s + 1) \cdots} = 0$$

Para um sistema do tipo 2,

$$K_a = \lim_{s \to 0} \frac{s^2 K(T_a s + 1)(T_b s + 1) \cdots}{s^2(T_1 s + 1)(T_2 s + 1) \cdots} = K$$

Para um sistema do tipo 3 ou maior,

$$K_a = \lim_{s \to 0} \frac{s^2 K(T_a s + 1)(T_b s + 1) \cdots}{s^N(T_1 s + 1)(T_2 s + 1) \cdots} = \infty \qquad (N \geq 3)$$

Portanto, o erro atuante estacionário para a entrada parábola unitária é

$e_{ss} = \infty$ para sistemas do tipo 0 ou tipo 1

$e_{ss} = \dfrac{1}{K}$ para sistemas do tipo 2

$e_{ss} = 0$ para sistemas do tipo 3 ou maior

Note que tanto sistemas do tipo 0 como do tipo 1 são incapazes de seguir uma entrada em parábola em regime estacionário. O sistema do tipo 2 com realimentação unitária pode seguir uma entrada em parábola com um sinal de erro atuante finito. A Fig. 7.3 mostra um exemplo da resposta de um sistema do tipo 2 com realimentação unitária para um entrada em parábola. O sistema do tipo 3 ou maior com realimentação unitária segue uma entrada em parábola com erro atuante nulo em regime estacionário.

Resumo. A Tabela 7.1 resume os erros estacionários para sistemas do tipo 0, tipo 1 e tipo 2 quando sujeitos a várias entradas. Os valores finitos para erros

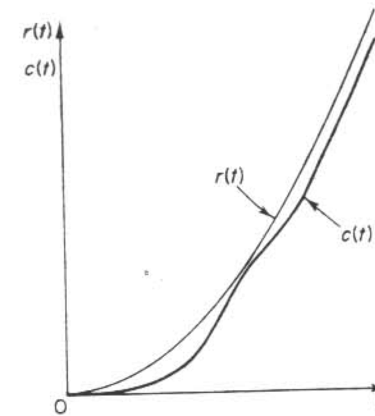


Fig. 7.3 Resposta de um sistema com realimentação unitária, do tipo 2, para uma entrada na parábola.

Tabela 7.1 Erro estacionário em termos do ganho K

	Entrada em degrau $r(t) = 1$	Entrada em rampa $r(t) = t$	Entrada de aceleração $r(t) = \frac{1}{2}t^2$
Sistema do tipo 0	$\frac{1}{1+K}$	∞	∞
Sistema do tipo 1	0	$\frac{1}{K}$	∞
Sistema do tipo 2	0	0	$\frac{1}{K}$

estacionários aparecem na linha diagonal. Acima da diagonal, os erros estacionários são infinitos; abaixo da diagonal eles são nulos.

Lembre-se de que os termos *erro de posição, erro de velocidade*, e *erro de aceleração* significam desvios em regime estacionário na posição de saída. Um erro de velocidade finita significa que depois de os transitórios desaparecerem, a entrada e a saída se movem com a mesma velocidade mas com uma diferença de posição finita.

Os coeficientes de erro K_p, K_v e K_a descrevem a habilidade de um sistema em reduzir ou eliminar erros estacionários. Portanto, eles são indicativos do desempenho em regime estacionário. Em geral é desejável aumentar os coeficientes de erro, enquanto se mantém a resposta transitória dentro de limites aceitáveis. Se há qualquer conflito entre o coeficiente de erro de velocidade estático e o coeficiente de erro de aceleração estático, então o último pode ser considerado menos importante do que o anterior. Nota-se que para melhorar o desempenho em regime estacionário, podemos aumentar o tipo do sistema adicionando uma integração ou integrações no ramo direto. Entretanto, isto introduz problemas adicionais de

estabilidade. O projeto de um sistema satisfatório com mais do que duas integrações no ramo direto é geralmente difícil.

Correlação entre a integral do erro na resposta a degrau e erro estacionário na resposta a rampa. Discutiremos a seguir a correlação entre a integral do erro na resposta a degrau unitário de um sistema e o erro estacionário na resposta a rampa unitária do mesmo sistema.

Em um sistema de controle com realimentação unitária, a área total abaixo da curva de erro $\int_0^\infty e(t)dt$ como resultado da resposta a degrau unitário fornece o erro estacionário na resposta a entrada rampa unitária. Para provar isto, vamos definir

$$\mathscr{L}[e(t)] = \int_0^\infty \varepsilon^{-st} e(t)\, dt = E(s)$$

Então

$$\lim_{s\to 0} \int_0^\infty \varepsilon^{-st} e(t)\, dt = \int_0^\infty e(t)\, dt = \lim_{s\to 0} E(s)$$

Note que

$$\frac{E(s)}{R(s)} = 1 - \frac{C(s)}{R(s)} = \frac{1}{1 + G(s)}$$

Portanto

$$\int_0^\infty e(t)\, dt = \lim_{s\to 0} \left[\frac{R(s)}{1 + G(s)}\right]$$

Para uma entrada degrau unitário,

$$\begin{aligned}\int_0^\infty e(t)\, dt &= \lim_{s\to 0} \left[\frac{1}{1 + G(s)} \frac{1}{s}\right] \\ &= \lim_{s\to 0} \frac{1}{sG(s)} \\ &= \frac{1}{K_v} \\ &= \text{erro estacionário na resposta a rampa unitária}\end{aligned}$$

Portanto em um sistema de controle com realimentação unitária, temos

$$\int_0^\infty e(t)\, dt = e_{ssr} \tag{7.1}$$

onde

$e(t)$ = erro na resposta a degrau unitário
e_{ssr} = erro estacionário na resposta a rampa unitária

Se e_{ssr} é nulo, então $e(t)$ deve mudar de sinal pelo menos uma vez. Isto significa que um sistema com erro de velocidade nulo (um sistema tendo $K_v = \infty$) apresentará pelo menos um sobre-sinal, quando o sistema é sujeito a uma entrada em degrau.

Exemplo 7.1 Considere o sistema visto na Fig. 7.4. Vamos obter a resposta a degrau unitário do sistema, esboçar a curva de resposta, calcular graficamente a integral do sinal de erro, $\int_0^\infty e(t)dt$, e verificar a Eq. (7.1). (O limite superior de integração deve, naturalmente, ser aproximado por algum tempo finito T, onde T é o tempo além do qual a saída praticamente permanece unitária.)

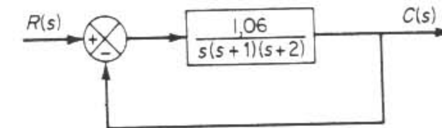


Fig. 7.4 Sistema de controle.

A função de transferência de malha-fechada para este sistema é dada por

$$\begin{aligned}\frac{C(s)}{R(s)} &= \frac{G(s)}{1 + G(s)} \\ &= \frac{1{,}06}{s(s+1)(s+2) + 1{,}06} \\ &= \frac{1{,}06}{(s + 2{,}33)(s + 0{,}33 + j0{,}58)(s + 0{,}33 - j0{,}58)}\end{aligned}$$

Para uma entrada degrau unitário,

$$R(s) = \frac{1}{s}$$

A resposta a degrau unitário é então obtida como

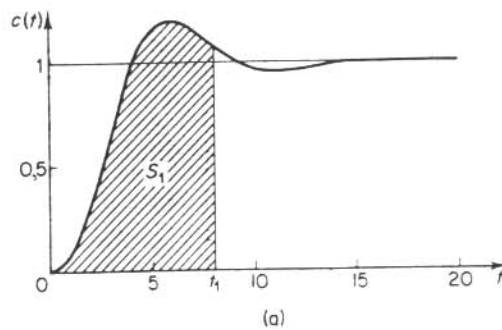
$$\begin{aligned}c(t) &= \mathscr{L}^{-1}[C(s)] \\ &= \mathscr{L}^{-1}\left[\frac{1{,}06}{s(s + 2{,}33)[(s + 0{,}33)^2 + 0{,}58^2]}\right]\end{aligned}$$

Referindo-se à fórmula,

$$\mathscr{L}^{-1}\left[\frac{1}{s(s+a)[(s+\zeta\omega_n)^2 + \omega_d^2]}\right] = \frac{1}{a\omega_n^2} - \frac{e^{-at}}{a[(a-\zeta\omega_n)^2 + \omega_d^2]}$$
$$+ \frac{e^{-\zeta\omega_n t}}{\omega_n[(a-\zeta\omega_n)^2 + \omega_d^2]}\left[\left(2\zeta - \frac{a}{\omega_n}\right)\cos\omega_d t + \frac{1}{\sqrt{1-\zeta^2}}\left(2\zeta^2 - \frac{a\zeta}{\omega_n} - 1\right)\text{sen}\,\omega_d t\right]$$

e observando que no presente sistema

$$a = 2{,}33, \quad \zeta = 0{,}5, \quad \omega_n = 0{,}667, \quad \omega_d = 0{,}58$$



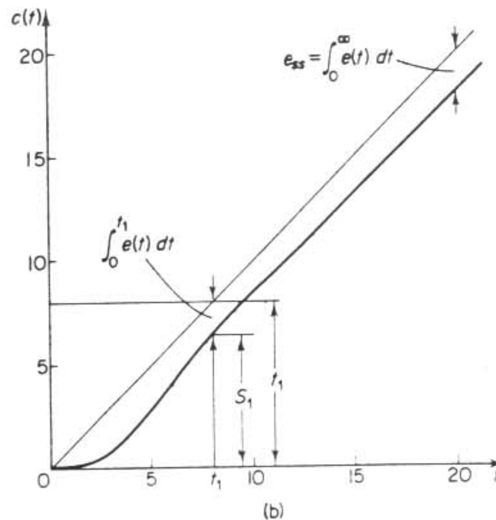


Fig. 7.5 (a) Resposta a degrau unitário do sistema visto na Fig. 7.4; (b) resposta a rampa unitária do sistema visto na Fig. 7.4.

obtemos $c(t)$ como segue:

$$c(t) = 1 - 0{,}103\,e^{-2{,}33t} - e^{-0{,}33t}(0{,}897\cos 0{,}58t + 0{,}933\text{sen}\,0{,}58t) \qquad (t \geq 0) \tag{7.2}$$

A Eq. (7.2) fornece a resposta a entrada degrau unitário; veja a Fig. 7.5(a).

Agora examinaremos a resposta do sistema a uma entrada rampa unitária. Quando todas as condições iniciais são nulas, a resposta a rampa unitária do sistema pode ser obtida integrando a curva de resposta a degrau unitário. A Fig. 7.5(b) mostra a curva de resposta a rampa unitária para o presente sistema. [Note que a área S_1 abaixo da curva de resposta a degrau unitário de $t = 0$ a $t = t_1$ dá o valor da saída em $t = t_1$ com uma entrada rampa unitária. Veja Figs. 7.5(a) e (b).]

No sistema presente, a resposta a degrau unitário se acomoda em aproximadamente $t = 20$. Portanto, a diferença entre 20 e a área abaixo da curva de resposta a degrau unitário de $t =$ 0 a $t = 20$ dá o erro estacionário na resposta a rampa unitária. Este valor do erro deve ser igual ao valor dado por $1/\lim_{s\to 0} sG(s)$.

Da curva de resposta da Fig. 7.5(a), obtemos os seguintes dados:

t	Área S abaixo da curva de resposta a degrau unitário de 0 a t	Área de erro $1 \times t - S$
0	0	0
2	0,16	1,84
4	1,50	2,50
6	3,80	2,20
8	6,10	1,90
10	8,15	1,85
12	10,10	1,90
14	12,06	1,94
16	14,07	1,93
18	16,08	1,92
20	18,08	1,92

A área de erro em regime estacionário é determinada como 1,92.

O coeficiente de erro de velocidade estático para este sistema é obtido a partir da função de transferência de malha-aberta como

$$K_v = \lim_{s\to 0} sG(s) = \frac{1{,}06}{2} = 0{,}53$$

Portanto o erro estacionário na resposta a rampa unitária é dado por

$$e_{ss} = \frac{1}{K_v} = 1{,}89$$

O valor obtido graficamente e o valor obtido analiticamente do erro estacionário na resposta a rampa unitária estão próximos um do outro. Portanto, verificamos a Eq. (7.1). (Note que se estes dois valores diferirem consideravelmente, então conferimos para verificar de onde vem a discrepância, achamos os erros e os corrigimos.)

7.2 COEFICIENTES DE ERRO DINÂMICO

Uma característica da definição de coeficientes de erro estático é que apenas um dos coeficientes assume um valor finito para um dado sistema. Os outros coeficientes são ou nulos ou infinitos. O erro estacionário obtido através dos coeficientes de erro estático ou é nulo, ou um valor finito não nulo, ou é infinito. Portanto, a variação do erro com o tempo não pode ser obtida através do uso de tais coeficientes. Os coeficientes de erro dinâmico apresentados nesta seção fornecem alguma informação sobre como o erro varia com o tempo; ou seja, se o erro estacionário do sistema com uma dada entrada cresce ou não proporcionalmente a t, t^2 etc.

Sistemas com erros dinâmicos diferentes mas com coeficientes de erro estático idênticos. Primeiro demonstraremos que dois sistemas com erros dinâmicos dife-

rentes podem ter coeficientes de erro estático idênticos. Considere os seguintes sistemas:

$$G_1(s) = \frac{10}{s(s+1)}, \qquad G_2(s) = \frac{10}{s(5s+1)}$$

Os coeficientes de erro estático são dados por

$$K_{p1} = \infty, \qquad K_{p2} = \infty$$
$$K_{v1} = 10, \qquad K_{v2} = 10$$
$$K_{a1} = 0, \qquad K_{a2} = 0$$

Portanto, os dois sistemas têm o mesmo erro estacionário para a mesma entrada em degrau. Comentários semelhantes se aplicam a erros estacionários para entradas em rampa e em parábola. Esta análise indica que é impossível estimar o erro dinâmico do sistema a partir dos coeficientes de erro estático.

Coeficientes de erro dinâmico. Agora introduziremos coeficientes de erro dinâmico para descrever erros dinâmicos. Limitaremos nossos sistemas para os de realimentação unitária. Dividindo o polinômio do numerador de $E(s)/R(s)$ pelo seu polinômio do denominador, $E(s)/R(s)$ pode ser expandido em uma série em potências ascendentes de s como segue:

$$\frac{E(s)}{R(s)} = \frac{1}{1+G(s)} = \frac{1}{k_1} + \frac{1}{k_2}s + \frac{1}{k_3}s^2 + \cdots$$

Os coeficientes $k_1, k_2, k_3, \ldots$ da série de potências são definidos como os *coeficientes de erro dinâmico*. Ou seja,

k_1 = coeficiente de erro de posição dinâmico
k_2 = coeficiente de erro de velocidade dinâmico
k_3 = coeficiente de erro de aceleração dinâmico

Em um dado sistema, os coeficientes de erro dinâmico estão relacionados com os coeficientes de erro estático. Considere o seguinte sistema do tipo 0 com realimentação unitária:

$$G(s) = \frac{K}{Ts+1}$$

Os coeficientes de erro de posição estático, os coeficientes de erro de velocidade estático, e os coeficientes de erro de aceleração estático são, respectivamente,

$$K_p = K$$
$$K_v = 0$$
$$K_a = 0$$

Como $E(s)/R(s)$ pode ser expandido como

$$\frac{E(s)}{R(s)} = \frac{1+Ts}{1+K+Ts} = \frac{1}{1+K} + \frac{TK}{(1+K)^2}s + \cdots$$

os coeficientes de erro dinâmico são dados em termos dos coeficientes de erro estático como segue:
O coeficiente de erro de posição dinâmico é

$$k_1 = 1 + K = 1 + K_p$$

O coeficiente de erro de velocidade dinâmico é

$$k_2 = \frac{(1+K)^2}{TK}$$

Como um outro exemplo, considere o sistema de controle com realimentação unitária com a seguinte função de transferência no ramo direto:

$$G(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s}$$

Os coeficientes de erro estático são dados por

$$K_p = \lim_{s\to 0} G(s) = \infty$$

$$K_v = \lim_{s\to 0} sG(s) = \frac{\omega_n}{2\zeta}$$

$$K_a = \lim_{s\to 0} s^2G(s) = 0$$

Como $E(s)/R(s)$ pode ser expandido como

$$\frac{E(s)}{R(s)} = \frac{s^2 + 2\zeta\omega_n s}{s^2 + 2\zeta\omega_n s + \omega_n^2} = \frac{\frac{2\zeta}{\omega_n}s + \frac{1}{\omega_n^2}s^2}{1 + \frac{2\zeta}{\omega_n}s + \frac{s^2}{\omega_n^2}} = \frac{2\zeta}{\omega_n}s + \left(\frac{1-4\zeta^2}{\omega_n^2}\right)s^2 + \cdots$$

o coeficiente de erro de velocidade dinâmico é igual ao coeficiente de erro de velocidade estático, ou seja,

$$k_2 = \frac{\omega_n}{2\zeta} = K_v$$

O coeficiente de erro de aceleração dinâmico é dado por

$$k_3 = \frac{\omega_n^2}{1-4\zeta^2}$$

Se uma análise similar é feita para sistemas de ordem superior, podemos

mostrar que para um sistema do tipo N, os coeficientes de erro dinâmico são dados por

$$k_{n+1} = \infty \qquad \text{para } n < N$$
$$k_{n+1} = \lim_{s\to 0} s^N G(s) \qquad \text{para } n = N$$

onde $n = 0, 1, 2, \ldots$ Os valores de k_{n+1} para $n > N$ são determinados pelos resultados da expansão de $E(s)/R(s)$ em torno da origem.

Vantagem dos coeficientes de erro dinâmico. Uma vantagem dos coeficientes de erro dinâmico se torna clara quando $E(s)$ é escrita da seguinte forma:

$$E(s) = \frac{1}{k_1}R(s) + \frac{1}{k_2}sR(s) + \frac{1}{k_3}s^2R(s) + \cdots$$

A região de convergência desta série é a vizinhança de $s = 0$. Isto corresponde a $t = \infty$ no domínio do tempo. A correspondente solução no tempo ou o erro estacionário é dado como segue, supondo todas as condições iniciais iguais a zero e desprezando impulsos em $t = 0$:

$$\lim_{t\to\infty} e(t) = \lim_{t\to\infty}\left[\frac{1}{k_1}r(t) + \frac{1}{k_2}\dot{r}(t) + \frac{1}{k_3}\ddot{r}(t) + \cdots\right]$$

O erro estacionário devido à função de entrada e suas derivadas pode então ser dado em termos dos coeficientes de erro dinâmico. Esta é uma vantagem dos coeficientes de erro dinâmico.

Exemplo 7.2 Determine os coeficientes de erro dinâmico do sistema de controle com realimentação unitária cuja função de transferência no ramo direto é dada por

$$G(s) = \frac{10}{s(s+1)}$$

Determine também o erro estacionário para a entrada definida por

$$r(t) = a_0 + a_1t + a_2t^2$$

Para o presente sistema,

$$\frac{E(s)}{R(s)} = \frac{1}{1+G(s)}$$
$$= \frac{s+s^2}{10+s+s^2}$$
$$= 0{,}1s + 0{,}09s^2 - 0{,}019s^3 + \cdots$$

ou

$$E(s) = 0{,}1sR(s) + 0{,}09s^2R(s) - 0{,}019s^3R(s) + \cdots$$

No domínio do tempo, o erro estacionário é dado por

$$\lim_{t\to\infty} e(t) = \lim_{t\to\infty}[0{,}1\dot{r}(t) + 0{,}09\ddot{r}(t) - 0{,}019\dddot{r}(t) + \cdots]$$

Os coeficientes de erro dinâmico são

$$k_1 = \infty$$
$$k_2 = \frac{1}{0{,}1} = 10$$
$$k_3 = \frac{1}{0{,}09} = 11{,}1$$

Como $r(t)$ é dado por

$$r(t) = a_0 + a_1t + a_2t^2$$

obtemos

$$\dot{r}(t) = a_1 + 2a_2t, \qquad \ddot{r}(t) = 2a_2, \qquad \dddot{r}(t) = 0$$

O erro estacionário é então

$$\lim_{t\to\infty} e(t) = \lim_{t\to\infty}[0{,}1(a_1 + 2a_2t) + 0{,}09(2a_2)]$$
$$= \lim_{t\to\infty}(0{,}1a_1 + 0{,}18a_2 + 0{,}2a_2t)$$

A menos que $a_2 = 0$, o erro estacionário se torna infinito.

Da análise apresentada pode-se ver que, se $E(s)/R(s)$ é expandido em torno da origem em uma série de potências, os coeficientes sucessivos da série indicam o erro dinâmico do sistema quando este é sujeito a uma entrada que varia lentamente. Os coeficientes de erro dinâmico fornecem uma maneira simples de se estimar o sinal de erro para entradas arbitrárias e o erro estacionário, sem a necessidade de se resolver o sistema de equações diferenciais.

Uma observação sobre os coeficientes de erro de velocidade dinâmico. É importante ressaltar que o coeficiente de erro de velocidade dinâmico k_2 pode ser estimado a partir da constante de tempo do sistema de primeira-ordem que aproxima a dada função de transferência de malha-fechada nas vizinhanças de $s = 0$.

Considere um sistema de controle com realimentação unitária com a seguinte função de transferência de malha-fechada:

$$\frac{C(s)}{R(s)} = \frac{1 + T_as + T_bs^2 + \cdots}{1 + T_1s + T_2s^2 + \cdots}$$

Nas vizinhanças de $s = 0$,

$$\lim_{s\to 0}\frac{C(s)}{R(s)} = \frac{1}{1 + (T_1 - T_a)s + \cdots}$$

O coeficiente de erro de velocidade dinâmico k_2 é então dado por

$$k_2 = \frac{1}{T_1 - T_a} \tag{7.3}$$

Isto pode ser verificado expandindo-se a função $E(s)/R(s)$ em torno da origem como segue:

$$\begin{aligned}\frac{E(s)}{R(s)} &= 1 - \frac{C(s)}{R(s)} \\ &= 1 - \frac{1}{1 + (T_1 - T_a)s} \\ &= 1 - 1 + (T_1 - T_a)s + \cdots \\ &= (T_1 - T_a)s + \cdots\end{aligned}$$

O coeficiente de erro de velocidade dinâmico é então dado pela Eq. (7.3).

Da análise que vimos, podemos concluir que se $C(s)/R(s)$ é aproximado por

$$\frac{C(s)}{R(s)} = \frac{1}{1 + T_{eq}s} \qquad \text{para } s \ll 1$$

então o coeficiente de erro de velocidade dinâmico k_2 é

$$k_2 = \frac{1}{T_{eq}}$$

Note que o coeficiente de erro de velocidade dinâmico k_2 obtido desta forma é o mesmo que o coeficiente de erro de velocidade estático. Como $C(s)/R(s)$ pode ser escrito como

$$\begin{aligned}\frac{C(s)}{R(s)} &= \frac{1}{1 + T_{eq}s} \\ &= \frac{\dfrac{1}{T_{eq}s}}{1 + \dfrac{1}{T_{eq}s}} \\ &= \frac{G(s)}{1 + G(s)}\end{aligned}$$

o coeficiente de erro de velocidade estático K_v é

$$\begin{aligned}K_v &= \lim_{s \to 0} sG(s) \\ &= \lim_{s \to 0} s\frac{1}{T_{eq}s} \\ &= \frac{1}{T_{eq}}\end{aligned}$$

7.3 CRITÉRIOS DE ERRO

No projeto de um sistema de controle, é importante que o sistema satisfaça as especificações de desempenho dadas. Como os sistemas de controle são dinâmicos, as especificações de desempenho podem ser dadas em termos do comportamento da resposta transitória para determinadas entradas, como entradas em degrau, entradas em rampa etc., ou em termos de um índice de desempenho.

Índices de desempenho. Um índice de desempenho é um número que indica a qualidade do desempenho do sistema. Um sistema de controle é considerado ótimo se os valores dos parâmetros são escolhidos de tal forma que o índice de desempenho escolhido é minimizado ou maximizado. Os valores ótimos dos parâmetros dependem diretamente do índice de desempenho selecionado.

Características necessárias para índices de desempenho. Um índice de desempenho deve oferecer seletividade; isto é, um ajuste ótimo dos parâmetros deve ser distinguido claramente de um ajuste não ótimo dos parâmetros. Adicionalmente, um índice de desempenho deve fornecer um único número positivo ou nulo, com o último sendo obtido se e apenas se a medida do desvio é identicamente nula. Para ser útil, um índice de desempenho deve ser uma função dos parâmetros do sistema e deve apresentar um máximo ou mínimo. Finalmente, para ser prático, um índice de desempenho deve ser computado com facilidade, analítica ou experimentalmente.

Índices de desempenho de erro. Em seguida, discutiremos vários critérios de erro em que os correspondentes índices de desempenho são integrais de alguma função ou função ponderada do desvio da saída do sistema real relativo à saída desejada. Como os valores das integrais podem ser obtidos como funções dos parâmetros do sistema, uma vez especificado o índice de desempenho, o sistema ótimo pode ser projetado ajustando-se os parâmetros para fornecer, por exemplo, o menor valor para a integral.

Vários índices de desempenho de erro têm sido propostos na literatura especializada. Discutiremos os quatro seguintes nesta seção.

$$\int_0^\infty e^2(t)\,dt, \qquad \int_0^\infty te^2(t)\,dt, \qquad \int_0^\infty |e(t)|\,dt, \qquad \int_0^\infty t\,|e(t)|\,dt$$

Considere um sistema de controle cuja saída desejada é $x(t)$ e cuja saída real é $y(t)$. Iremos definir o erro $e(t)$ como

$$e(t) = x(t) - y(t)$$

Note que a menos que

$$\lim_{t \to \infty} e(t) = 0$$

os índices de desempenho tendem a infinito. Se $\lim_{t \to \infty} e(t)$ não tende a zero, podemos definir

$$e(t) = y(\infty) - y(t)$$

Com esta definição do erro, os índices de desempenho resultarão em números finitos.

Critério da integral do erro quadrático. De acordo com o critério da integral do erro quadrático (ISE, *integral square-error*), a qualidade do desempenho do sistema é avaliada pela seguinte integral:

$$\int_0^\infty e^2(t)\,dt$$

onde o limite superior ∞ pode ser substituído por T que é escolhido grande o suficiente para que $e(t)$ seja desprezível para $T < t$. O sistema ótimo é aquele que minimiza esta integral. Este índice de desempenho tem sido usado freqüentemente, tanto para entradas determinísticas (como entradas em degrau) quanto para entradas aleatórias, por causa da facilidade de se computar a integral tanto analítica quanto experimentalmente. A Fig. 7.6 mostra $x(t)$, $y(t)$, $e(t)$, $e^2(t)$, e $\int e^2(t)dt$ quando a saída desejada $x(t)$ é um degrau unitário. A integral de $e^2(t)$ de 0 a T é a área total abaixo da curva $e^2(t)$.

Uma característica deste índice de desempenho é que ele dá grande peso para erros grandes e pequeno peso para erros pequenos. Este critério não é muito seletivo pois, para o seguinte sistema de segunda-ordem,

$$\frac{C(s)}{R(s)} = \frac{1}{s^2 + 2\zeta s + 1}$$

uma variação em ζ de 0,5 a 0,7 não causa muita variação no valor da integral.

Um sistema projetado usando este critério tende a apresentar uma rápida diminuição em um erro inicial grande. Portanto a resposta é rápida e oscilatória. Desta forma, o sistema tem baixa estabilidade relativa.

Note, entretanto, que o critério de integral do erro quadrático freqüentemente

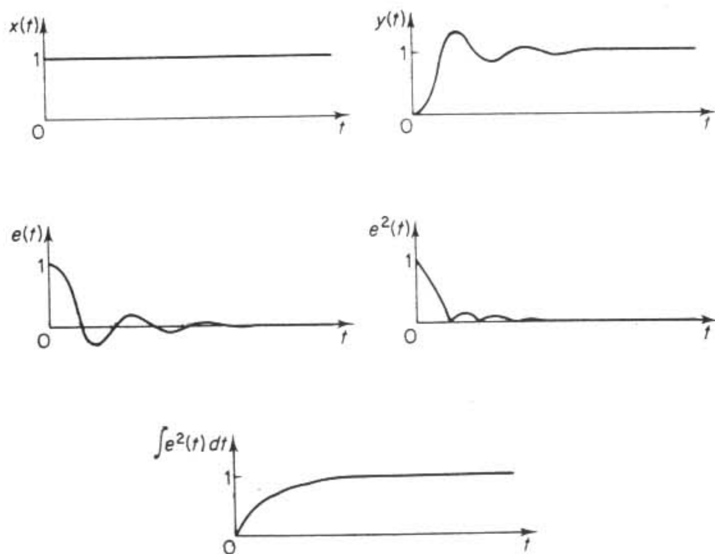


Fig. 7.6 Curvas mostrando a saída desejada $x(t)$, a saída real $y(t)$, o erro $e(t)$, o erro quadrático $e^2(t)$, e a integral do erro quadrático $\int e^2(t)\,dt$ em função de t.

é de significância prática porque a minimização do índice de desempenho resulta na minimização do consumo de potência para alguns sistemas, como por exemplo, sistemas de naves espaciais.

Critério da integral do erro quadrático multiplicado pelo tempo. O índice de desempenho baseado no critério da integral do erro quadrático multiplicado pelo tempo (ITSE, *integral-of-time-multiplied square-error*) é

$$\int_0^\infty t e^2(t)\,dt$$

O sistema ótimo é aquele que minimiza esta integral.

Uma característica deste critério é que, na resposta a degrau unitário do sistema, um erro inicial grande é ponderado com peso baixo, enquanto que erros que ocorrem mais tarde na resposta transitória são bastante penalizados. Este critério tem melhor seletividade do que o critério da integral do erro quadrático.

Critério da integral do erro absoluto. O índice de desempenho definido pelo critério da integral do erro absoluto (IAE, *integral absolute-error*) é

$$\int_0^\infty |e(t)|\,dt$$

Este é um dos índices de desempenho de mais fácil aplicação. Se este critério é usado, tanto sistemas altamente subamortecidos como altamente sobreamortecidos não podem ser feitos ótimos. Um sistema ótimo baseado neste critério é um sistema que tem um amortecimento razoável e uma característica de resposta transitória satisfatória; entretanto, a seletividade deste índice de desempenho não é muito boa. Embora este índice de desempenho não possa ser facilmente computado analiticamente, ele é particularmente conveniente para estudos com computador analógico. A Fig. 7.7 mostra a curva de $e(t)$ *versus* t e a curva $|e(t)|$ *versus* t.

Note que a minimização da integral do erro absoluto está diretamente relacionada com a minimização do consumo de combustível de sistemas aeroespaciais.

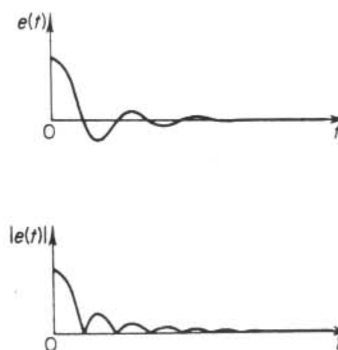


Fig. 7.7 Curvas mostrando $e(t)$ em função de t e $|e(t)|$ em função de t.

Critério da integral do erro absoluto multiplicado pelo tempo. De acordo com este critério, o critério ITAE (*integral-of-time-multiplied absolute-error*), o sistema ótimo é aquele que minimiza o seguinte índice de desempenho:

$$\int_0^\infty t\,|e(t)|\,dt$$

Como nos critérios precedentes, um erro inicial grande em uma resposta a degrau unitário é ponderado com pequeno peso, e erros que ocorrem mais tarde na resposta transitória são bastante penalizados.

A característica de um sistema projetado que usa este critério é que o sobre-sinal na resposta transitória é pequeno e oscilações são bem amortecidas. Este critério possui boa seletividade e constitui uma melhora sobre o critério de integral do erro absoluto. Entretanto, ele é difícil de ser computado analiticamente, embora possa ser medido experimentalmente com facilidade.

Comparação dos vários critérios de erro. A Fig. 7.8 mostra várias curvas de desempenho de erro. O sistema considerado é

$$\frac{C(s)}{R(s)} = \frac{1}{s^2 + 2\zeta s + 1}$$

A inspeção destas curvas mostra que o critério da integral do erro quadrático não tem boa seletividade porque a curva é um tanto achatada perto do ponto onde o índice de desempenho é mínimo. A adição do quadrado da taxa de erro ao integrando do índice de desempenho, ou

$$\int_0^\infty [e^2(t) + \dot{e}^2(t)]\,dt$$

leva o valor ótimo de ζ para 0,707, mas a seletividade é novamente ruim. Da Fig. 7.8 vemos claramente que os outros índices de desempenho têm boa seletividade.

As curvas na Fig. 7.8 indicam que, no sistema de segunda-ordem em consideração, $\zeta = 0.7$ corresponde ao valor ótimo ou quase-ótimo com relação a cada um dos índices de desempenho utilizados. Quando o coeficiente de amortecimento ζ é 0.7, o sistema de segunda-ordem resulta com uma resposta rápida para entrada em degrau com aproximadamente 5% de sobre-sinal.

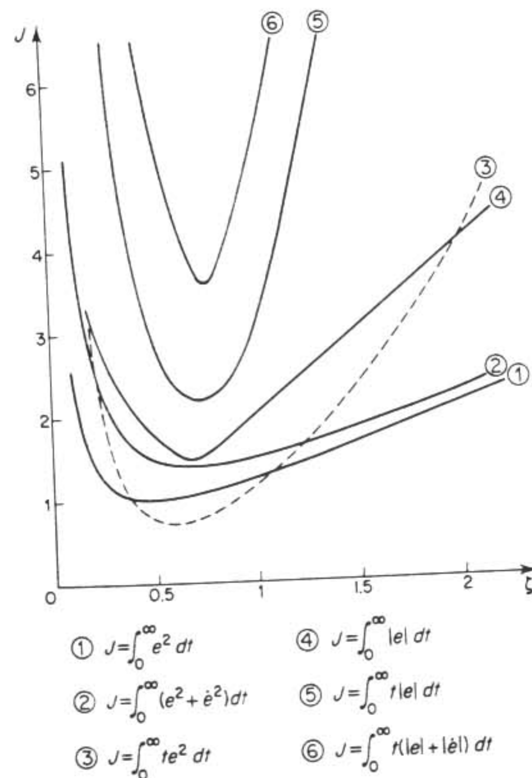


Fig. 7.8 Curvas de desempenho de erro.

Aplicação do critério ITAE para sistemas de ordem n.* O critério ITAE foi utilizado para o seguinte sistema de ordem n:

$$\frac{C(s)}{R(s)} = \frac{a_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}$$

e os coeficientes que minimizam

$$\int_0^\infty t\,|e|\,dt$$

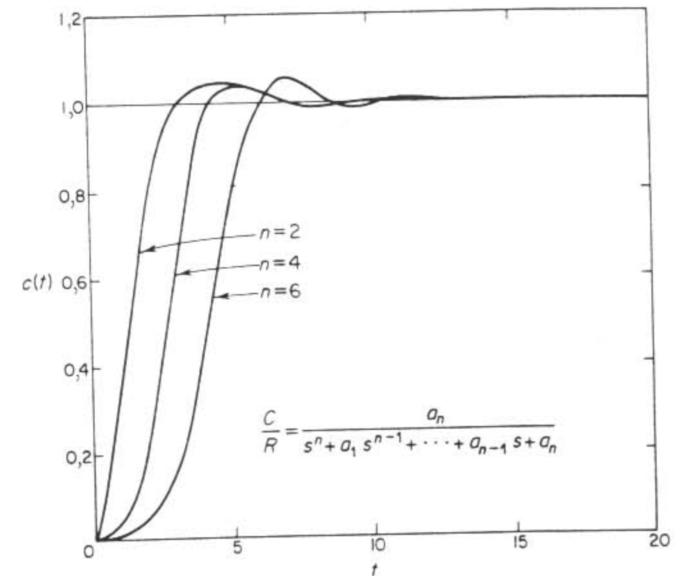


Fig. 7.9 Curvas de resposta a degrau de sistemas com funções de transferência ótimas baseadas no critério ITAE.

*Referência G-5.

foram determinados. Claramente, esta função de transferência resulta em um erro estacionário nulo para entradas em degrau.

A Tabela 7.2 mostra a forma ótima da função de transferência de malha-fechada baseada neste critério ITAE. A Fig. 7.9 mostra as curvas de resposta a degrau dos sistemas ótimos.

Tabela 7.2 Forma ótima da função de transferência de malha-fechada baseada no critério ITAE (sistemas de erro nulo para degraus)

$\frac{C(s)}{R(s)} = \frac{a_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}, \quad a_n = \omega_n^n$
$s + \omega_n$
$s^2 + 1{,}4\omega_n s + \omega_n^2$
$s^3 + 1{,}75\omega_n s^2 + 2{,}15\omega_n^2 s + \omega_n^3$
$s^4 + 2{,}1\omega_n s^3 + 3{,}4\omega_n^2 s^2 + 2{,}7\omega_n^3 s + \omega_n^4$
$s^5 + 2{,}8\omega_n s^4 + 5{,}0\omega_n^2 s^3 + 5{,}5\omega_n^3 s^2 + 3{,}4\omega_n^4 s + \omega_n^5$
$s^6 + 3{,}25\omega_n s^5 + 6.60\omega_n^2 s^4 + 8{,}60\omega_n^3 s^3 + 7{,}45\omega_n^4 s^2 + 3{,}95\omega_n^5 s + \omega_n^6$

7.4 INTRODUÇÃO À OTIMIZAÇÃO DE SISTEMAS

Nesta seção, resolveremos problemas simples de otimização em que minimizamos certos índices de desempenho de erro. Introduziremos a técnica direta e a técnica da transformada de Laplace para resolver este problema. Apresentaremos uma técnica alternativa (Liapunov) para a solução de problemas de otimização similares no Cap. 16.

Computação direta dos índices de desempenho "ISE". Considere o sistema de controle visto na Fig. 7.10. Suponha que uma entrada degrau unitário $r(t)$ é aplicada ao sistema em $t = 0$. Supomos que o sistema está inicialmente em repouso. O problema é determinar o valor ótimo de $\zeta > 0$ de tal forma que

$$J = \int_0^\infty x^2(t)\,dt \tag{7.4}$$

seja mínimo, onde $x(t)$ é o sinal de erro.

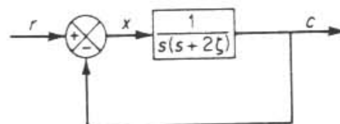


Fig. 7.10 Sistema de controle.

Da Fig. 7.10, a equação do sistema é obtida como segue:

$$\ddot{c} + 2\zeta\dot{c} + c = r$$

Em termos do sinal de erro $x = r - c$,

$$\ddot{x} + 2\zeta\dot{x} + x = 0 \qquad (t > 0) \tag{7.5}$$

A solução da Eq. (7.5) pode ser escrita como

$$x(t) = k_1 e^{\lambda_1 t} + k_2 e^{\lambda_2 t} \tag{7.6}$$

onde

$$\lambda_1 = -\zeta + \sqrt{\zeta^2 - 1}, \qquad \lambda_2 = -\zeta - \sqrt{\zeta^2 - 1}$$

k_1 e k_2 são determinados das condições iniciais.

Substituindo a Eq. (7.6) na Eq. (7.4) e observando que λ_1 e λ_2 têm partes reais negativas, obtemos

$$\begin{aligned} J &= \int_0^\infty x^2(t)\,dt \\ &= \int_0^\infty (k_1^2 e^{2\lambda_1 t} + 2k_1k_2 e^{(\lambda_1+\lambda_2)t} + k_2^2 e^{2\lambda_2 t})\,dt \\ &= -\frac{k_1^2}{2\lambda_1} - \frac{2k_1k_2}{\lambda_1 + \lambda_2} - \frac{k_2^2}{2\lambda_2} \end{aligned}$$

Como $\lambda_1 + \lambda_2 = -2\zeta$, $\lambda_1\lambda_2 = 1$, podemos simplificar J como segue:

$$J = -\frac{k_1^2\lambda_2 + k_2^2\lambda_1}{2} + \frac{k_1k_2}{\zeta} \tag{7.7}$$

As constantes k_1 e k_2 são determinadas das condições iniciais como segue:

$$1 = k_1 + k_2, \qquad 0 = k_1\lambda_1 + k_2\lambda_2$$

ou

$$k_1 = \frac{\lambda_2}{\lambda_2 - \lambda_1}, \qquad k_2 = -\frac{\lambda_1}{\lambda_2 - \lambda_1}$$

Substituindo os valores de k_1 e k_2 na Eq. (7.7), obtemos

$$\begin{aligned} J &= -\frac{1}{2(\lambda_2 - \lambda_1)^2}\left(\lambda_2^3 + \lambda_1^3 + \frac{2\lambda_1\lambda_2}{\zeta}\right) \\ &= -\frac{1}{2(\lambda_2 - \lambda_1)^2}\left\{(\lambda_1 + \lambda_2)[(\lambda_1 + \lambda_2)^2 - 3\lambda_1\lambda_2] + \frac{2\lambda_1\lambda_2}{\zeta}\right\} \\ &= \zeta + \frac{1}{4\zeta} \end{aligned}$$

O valor ótimo de ζ pode ser obtido diferenciando J com relação a ζ, equacionando $dJ/d\zeta$ a zero, e resolvendo para ζ. Como

$$\frac{dJ}{d\zeta} = 1 - \frac{1}{4\zeta^2} = 0$$

o valor ótimo de $\zeta > 0$ é

$$\zeta = 0{,}5$$

(Claramente, $\zeta = 0{,}5$ corresponde a um mínimo pois $d^2J/d\zeta^2 > 0$.) O valor mínimo de J é

$$\text{mín. } J = 0{,}5 + 0{,}5 = 1$$

Teorema da integração real. A seguir computaremos $\int_0^\infty f(t)dt$ usando o método da transformada de Laplace. Fazendo isto, usaremos o seguinte teorema da integração real:

Teorema da integração real: Se $f(t)$ é de ordem exponencial, então a transformada de Laplace de $\int_0^t f(t)dt$ existe e é

$$\mathscr{L}\left[\int_0^t f(t)\,dt\right] = \frac{F(s)}{s}$$

onde

$$F(s) = \mathscr{L}[f(t)]$$

Para provar este teorema, usamos integração por partes.

$$\begin{aligned}\mathscr{L}\left[\int_0^t f(t)\,dt\right] &= \int_0^\infty \left[\int_0^t f(t)\,dt\right]e^{-st}\,dt \\ &= -\frac{e^{-st}}{s}\int_0^t f(t)\,dt\Big|_0^\infty + \frac{1}{s}\int_0^\infty f(t)e^{-st}\,dt \\ &= \lim_{t\to\infty}\left[-\frac{e^{-st}}{s}\int_0^t f(t)\,dt\right] + \frac{F(s)}{s}\end{aligned}$$

Como $f(t)$ é de ordem exponencial, quando a parte real de s é suficientemente grande, obtemos

$$\lim_{t\to\infty}\frac{e^{-st}}{s}\int_0^t f(t)\,dt \longrightarrow 0$$

Portanto

$$\mathscr{L}\left[\int_0^t f(t)\,dt\right] = \frac{F(s)}{s}$$

Isto prova o teorema.

Técnica da transformada de Laplace para a computação de índices de desempenho "ISE". Usando os teoremas da integração e do valor final, obtemos

$$\int_0^\infty f(t)\,dt = \lim_{t\to\infty}\int_0^t f(t)\,dt = \lim_{s\to 0} s\frac{F(s)}{s} = \lim_{s\to 0} F(s)$$

Esta é a equação básica para calcular o índice de desempenho ISE. No problema presente,

$$f(t) = x^2(t)$$

Portanto

$$\lim_{t\to\infty}\int_0^t x^2(t)\,dt = \lim_{s\to 0} F(s)$$

onde

$$F(s) = \mathscr{L}[x^2(t)]$$

Considere outra vez o sistema de controle visto na Fig. 7.10. Para este sistema,

$$\frac{X(s)}{R(s)} = \frac{s^2 + 2\zeta s}{s^2 + 2\zeta s + 1}$$

Portanto, para a entrada degrau unitário,

$$X(s) = \frac{s^2 + 2\zeta s}{s^2 + 2\zeta s + 1}\,\frac{1}{s} = \frac{s + \zeta}{s^2 + 2\zeta s + 1} + \frac{\zeta}{s^2 + 2\zeta s + 1}$$

A transformada inversa de Laplace é

$$x(t) = e^{-\zeta t}\left(\cos\sqrt{1-\zeta^2}\,t + \frac{\zeta}{\sqrt{1-\zeta^2}}\,\text{sen}\sqrt{1-\zeta^2}\,t\right) \qquad (t \geq 0)$$

Portanto

$$x^2(t) = e^{-2\zeta t}\left[\frac{1}{2(1-\zeta^2)} + \frac{1}{2}\,\frac{1-2\zeta^2}{1-\zeta^2}\cos 2\sqrt{1-\zeta^2}\,t + \frac{\zeta}{\sqrt{1-\zeta^2}}\,\text{sen}\,2\sqrt{1-\zeta^2}\,t\right]$$

Então,

$$F(s) = \mathscr{L}[x^2(t)] = \frac{1}{2(1-\zeta^2)}\,\frac{1}{s+2\zeta} + \frac{1-2\zeta^2}{2(1-\zeta^2)}\,\frac{s+2\zeta}{(s+2\zeta)^2 + 4(1-\zeta^2)} + \frac{\zeta}{\sqrt{1-\zeta^2}}\,\frac{2\sqrt{1-\zeta^2}}{(s+2\zeta)^2 + 4(1-\zeta^2)} \qquad (7.8)$$

Tomando o limite quando s tende a zero, resulta

$$\lim_{s\to 0} F(s) = \frac{1}{2(1-\zeta^2)}\frac{1}{2\zeta} + \frac{1-2\zeta^2}{2(1-\zeta^2)}\frac{\zeta}{2} + \frac{\zeta}{2}$$
$$= \zeta + \frac{1}{4\zeta}$$

Portanto

$$J = \int_0^\infty x^2(t)\,dt$$
$$= \zeta + \frac{1}{4\zeta} \qquad (7.9)$$

Este valor de J, convém mencionar, é o mesmo que foi obtido pelo método direto, e portanto o valor ótimo de ζ que minimiza J é 0,5.

Teorema da diferenciação complexa. Ao calcular o índice de desempenho $\int_0^\infty tf(t)dt$, achamos conveniente usar o seguinte teorema da diferenciação complexa:

Teorema da diferenciação complexa: Se $f(t)$ é transformável segundo Laplace, então, exceto em pontos singulares de $F(s)$,

$$\mathscr{L}[tf(t)] = -\frac{d}{ds}F(s)$$

onde

$$\mathscr{L}[f(t)] = F(s)$$

Para provar isto, note que

$$\frac{d}{ds}F(s) = \frac{d}{ds}\int_0^\infty f(t)e^{-st}\,dt$$
$$= -\int_0^\infty tf(t)e^{-st}\,dt$$

Portanto está provado o teorema.

Técnica da transformada de Laplace para a computação de índices de desempenho "ITSE". Considere a seguir o sistema visto na Fig. 7.10, com o seguinte índice de desempenho:

$$J = \int_0^\infty tx^2(t)\,dt$$

Novamente supomos que o sistema está em repouso antes que a entrada degrau unitário seja aplicada. Achemos o valor ótimo de ζ que minimiza o índice de desempenho.

Usando os teoremas de integração real e diferenciação complexa, temos

$$\mathscr{L}[tx^2(t)] = -\frac{d}{ds}F(s)$$

onde

$$F(s) = \mathscr{L}[x^2(t)]$$

No problema presente, $F(s)$ é dado pela Eq. (7.8).

$$\mathscr{L}[tx^2(t)] = -\frac{d}{ds}F(s)$$
$$= \frac{1}{2(1-\zeta^2)}\frac{1}{(s+2\zeta)^2} + \frac{1-2\zeta^2}{2(1-\zeta^2)}\frac{(s+2\zeta)^2-4(1-\zeta^2)}{[(s+2\zeta)^2+4(1-\zeta^2)]^2}$$
$$+ \frac{4\zeta(s+2\zeta)}{[(s+2\zeta)^2+4(1-\zeta^2)]^2}$$

Usando o teorema do valor final,

$$\int_0^\infty tx^2(t)\,dt = \lim_{s\to 0}\left[-\frac{d}{ds}F(s)\right]$$

Tomando o limite de $-dF(s)/ds$ quando s tende a zero, obtemos

$$\lim_{s\to 0}\left[-\frac{d}{ds}F(s)\right] = \frac{1}{2(1-\zeta^2)}\frac{1}{4\zeta^2} + \frac{1-2\zeta^2}{2(1-\zeta^2)}\frac{2\zeta^2-1}{4} + \frac{\zeta^2}{2}$$
$$= \zeta^2 + \frac{1}{8\zeta^2}$$

Portanto

$$J = \int_0^\infty tx^2(t)\,dt$$
$$= \zeta^2 + \frac{1}{8\zeta^2} \qquad (7.10)$$

Para se obter o valor ótimo de ζ que minimiza J, diferenciamos J com relação a ζ:

$$\frac{dJ}{d\zeta} = 2\zeta - \frac{1}{4\zeta^3} = 0$$

que resulta no seguinte:

$$\zeta = 0{,}595$$

Claramente, $d^2J/d\zeta^2 > 0$. O valor ótimo de ζ é 0,595, e o valor mínimo de J é

$$\text{mín. } J = 0{,}595^2 + \frac{1}{8\times 0{,}595^2} = 0{,}71$$

PROBLEMAS ILUSTRATIVOS E SOLUÇÕES

Problema A.7.1 Determine o erro estacionário na saída de um sistema de controle linear com realimentação unitária quando a entrada é dada por

$$r(t) = r_0 + r_1 t + \frac{r_2}{2}t^2$$

Solução. O erro estacionário é a soma dos erros devidos a cada componente de entrada. Portanto, o erro estacionário é determinado como

$$e_{ss} = \frac{r_0}{1 + K_p} + \frac{r_1}{K_v} + \frac{r_2}{K_a}$$

O erro estacionário se torna infinito a menos que o sistema seja de ordem 2 ou maior.

Problema A.7.2 Considere o sistema de controle com realimentação unitária com função de transferência do ramo direto $G(s)$. Suponha que a função de transferência de malha-fechada pode ser escrita como

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1+G(s)} = \frac{(T_a s + 1)(T_b s + 1) \cdots (T_m s + 1)}{(T_1 s + 1)(T_2 s + 1) \cdots (T_n s + 1)}$$

Mostre que

$$\int_0^\infty e(t)dt = (T_1 + T_2 + \cdots + T_n) - (T_a + T_b + \cdots + T_m)$$

onde $e(t)$ é o erro na resposta a degrau unitário. Mostre também que

$$\frac{1}{\lim_{s\to 0} sG(s)} = (T_1 + T_2 + \cdots + T_n) - (T_a + T_b + \cdots + T_m)$$

Solução. Vamos definir

$$(T_a s + 1)(T_b s + 1) \cdots (T_m s + 1) = P(s)$$

e

$$(T_1 s + 1)(T_2 s + 1) \cdots (T_n s + 1) = Q(s)$$

Então

$$\frac{C(s)}{R(s)} = \frac{P(s)}{Q(s)}$$

e

$$E(s) = \frac{Q(s) - P(s)}{Q(s)} R(s)$$

Para uma entrada degrau unitário, $R(s) = 1/s$. Portanto,

$$E(s) = \frac{Q(s) - P(s)}{sQ(s)}$$

E usando o fato de que

$$\int_0^\infty e(t)\,dt = \lim_{s\to 0} s\frac{E(s)}{s} = \lim_{s\to 0} E(s)$$

obtemos

$$\begin{aligned}\int_0^\infty e(t)\,dt &= \lim_{s\to 0} \frac{Q(s) - P(s)}{sQ(s)} \\ &= \lim_{s\to 0} \frac{Q'(s) - P'(s)}{Q(s) + sQ'(s)} \\ &= \lim_{s\to 0} [Q'(s) - P'(s)]\end{aligned}$$

Como

$$\lim_{s\to 0} P'(s) = T_a + T_b + \cdots + T_m$$

$$\lim_{s\to 0} Q'(s) = T_1 + T_2 + \cdots + T_n$$

temos

$$\int_0^\infty e(t)\,dt = (T_1 + T_2 + \cdots + T_n) - (T_a + T_b + \cdots + T_m)$$

Observe que, como

$$\int_0^\infty e(t)\,dt = \frac{1}{K_v} = \frac{1}{\lim_{s\to 0} sG(s)}$$

imediatamente obtemos a seguinte relação:

$$\frac{1}{\lim_{s\to 0} sG(s)} = (T_1 + T_2 + \cdots + T_n) - (T_a + T_b + \cdots + T_m)$$

Observe que zeros irão melhorar K_v. Pólos perto da origem causam coeficientes de erro de velocidade baixos a menos que haja zeros na vizinhança.

Problema A.7.3 Compare os erros dos sistemas de controle com realimentação unitária tendo as seguintes funções de transferência de ramo direto:

$$G_1(s) = \frac{10}{s(s+1)}, \qquad G_2(s) = \frac{10}{s(2s+1)}$$

Supõe-se que a entrada é

$$r(t) = 5 + 2t + 3t^2$$

Solução. Como

$$\frac{E_1(s)}{R(s)} = \frac{1}{1 + G_1(s)} = \frac{s + s^2}{10 + s + s^2}$$
$$= 0{,}1s + 0{,}09s^2 - 0{,}019s^3 + \cdots$$

e

$$\frac{E_2(s)}{R(s)} = \frac{1}{1 + G_2(s)} = \frac{s + 2s^2}{10 + s + 2s^2}$$
$$= 0{,}1s + 0{,}19s^2 - 0{,}039s^3 + \cdots$$

as funções temporais de erro são dadas por

$$\lim_{t\to\infty} e_1(t) = \lim_{t\to\infty} [0{,}1(2 + 6t) + 0{,}09 \times 6]$$

e

$$\lim_{t\to\infty} e_2(t) = \lim_{t\to\infty} [0{,}1(2 + 6t) + 0{,}19 \times 6]$$

Segue-se que

$$\lim_{t\to\infty} e_1(t) < \lim_{t\to\infty} e_2(t)$$

Problema A.7.4 Considere o sistema cuja função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{1}{s^2 + 2\zeta s + 1} \qquad (0 < \zeta < 1)$$

Calcule $\int_0^\infty c^2(t)dt$ para uma entrada impulso unitário.

Solução. Para uma entrada impulso unitário, $R(s) = 1$. Portanto,

$$C(s) = \frac{1}{s^2 + 2\zeta s + 1} = \frac{1}{\sqrt{1-\zeta^2}}\frac{\sqrt{1-\zeta^2}}{(s+\zeta)^2 + 1 - \zeta^2}$$

A transformada inversa de Laplace de $C(s)$ é

$$c(t) = \frac{1}{\sqrt{1-\zeta^2}}e^{-\zeta t}\operatorname{sen}\sqrt{1-\zeta^2}\,t \qquad (t \geq 0)$$

Como $c^2(t)$ é

$$c^2(t) = \frac{1}{1-\zeta^2}e^{-2\zeta t}\operatorname{sen}^2\sqrt{1-\zeta^2}\,t$$
$$= \frac{1}{1-\zeta^2}e^{-2\zeta t}\frac{1}{2}(1 - \cos 2\sqrt{1-\zeta^2}\,t)$$

obtemos

$$\mathscr{L}[c^2(t)] = \frac{1}{2(1-\zeta^2)}\left[\frac{1}{s+2\zeta} - \frac{s+2\zeta}{(s+2\zeta)^2 + 4(1-\zeta^2)}\right]$$

Então

$$\int_0^\infty c^2(t)\,dt = \lim_{s\to 0} \mathscr{L}[c^2(t)]$$
$$= \lim_{s\to 0} \frac{1}{2(1-\zeta^2)}\left[\frac{1}{s+2\zeta} - \frac{s+2\zeta}{(s+2\zeta)^2 + 4(1-\zeta^2)}\right]$$
$$= \frac{1}{4\zeta}$$

Problema A.7.5 Considere o seguinte sistema:

$$\ddot{x} + 2\zeta\dot{x} + x = 0 \qquad (0 < \zeta \leq 0{,}9) \tag{7.11}$$

Determine o valor de ζ que minimiza o seguinte índice de desempenho J:

$$J = \int_0^\infty [x^2(t) + \dot{x}^2(t)]\,dt$$

Solução. A solução da Eq. (7.11) é

$$x(t) = k_1e^{\lambda_1 t} + k_2e^{\lambda_2 t}$$

onde

$$\lambda_1 = -\zeta + \sqrt{\zeta^2 - 1}, \qquad \lambda_2 = -\zeta - \sqrt{\zeta^2 - 1}$$

k_1 e k_2 são determinados a partir das condições iniciais. Portanto,

$$\dot{x}(t) = k_1\lambda_1e^{\lambda_1 t} + k_2\lambda_2e^{\lambda_2 t}$$

Então $x^2(t)$ e $\dot{x}^2(t)$ são obtidos como

$$x^2(t) = k_1^2e^{2\lambda_1 t} + 2k_1k_2e^{(\lambda_1+\lambda_2)t} + k_2^2e^{2\lambda_2 t}$$

e

$$\dot{x}^2(t) = k_1^2\lambda_1^2e^{2\lambda_1 t} + 2k_1k_2\lambda_1\lambda_2e^{(\lambda_1+\lambda_2)t} + k_2^2\lambda_2^2e^{2\lambda_2 t}$$

O índice de desempenho então se torna

$$J = \int_0^\infty [x^2(t) + \dot{x}^2(t)]\,dt$$
$$= \int_0^\infty [k_1^2(1 + \lambda_1^2)e^{2\lambda_1 t} + 2k_1k_2(1 + \lambda_1\lambda_2)e^{(\lambda_1+\lambda_2)t} + k_2^2(1 + \lambda_2^2)e^{2\lambda_2 t}]\,dt$$

Fazendo a integração e calculando em $t = \infty$ e $t = 0$, e observando que as partes reais de λ_1 e λ_2 são negativas, obtemos

$$J = -\left[\frac{k_1^2(1+\lambda_1^2)}{2\lambda_1} + \frac{2k_1k_2(1+\lambda_1\lambda_2)}{\lambda_1+\lambda_2} + \frac{k_2^2(1+\lambda_2^2)}{2\lambda_2}\right]$$

Como

$$\lambda_1 + \lambda_2 = -2\zeta, \qquad \lambda_1\lambda_2 = 1$$

obtemos

$$J = \zeta(k_1^2 + k_2^2) - \frac{2k_1k_2}{\zeta}$$
$$= \zeta(k_1 + k_2)^2 - 2k_1k_2\left(\zeta - \frac{1}{\zeta}\right) \tag{7.12}$$

Das condições iniciais,

$$x(0) = c_1 = k_1 + k_2, \qquad \dot{x}(0) = c_2 = k_1\lambda_1 + k_2\lambda_2$$

Como $0 < \zeta \leq 0,9$, $\lambda_1 \neq \lambda_2$, k_1 e k_2 são obtidos como

$$k_1 = \frac{c_1\lambda_2 - c_2}{\lambda_2 - \lambda_1}, \qquad k_2 = \frac{c_2 - \lambda_1 c_1}{\lambda_2 - \lambda_1}$$

Substituindo estes valores de k_1 e k_2 na Eq. (7.12), obtemos

$$J = \zeta c_1^2 + c_1c_2 + \frac{c_1^2 + c_2^2}{2\zeta} \tag{7.13}$$

Para achar o valor mínimo de J, primeiro diferencie J com relação a ζ e então iguale o resultado a zero.

$$\frac{\partial J}{\partial \zeta} = c_1^2 - \frac{c_1^2 + c_2^2}{2\zeta^2} = 0$$

Resolvendo esta equação,

$$\zeta = \sqrt{\frac{c_1^2 + c_2^2}{2c_1^2}} \tag{7.14}$$

Como $\partial^2 J/\partial\zeta^2 > 0$, a Eq. (7.14) fornece o valor ótimo de ζ, contanto que

$$\sqrt{\frac{c_1^2 + c_2^2}{2c_1^2}} \leq 0,9$$

ou

$$c_2^2 \leq 0,62c_1^2$$

No caso especial, quando $c_2 = 0$,

$$\zeta = \frac{1}{\sqrt{2}}$$

é o valor ótimo de ζ. Portanto, se $c_1 = 1$ e $c_2 = 0$, então

$$\text{mín. } J = \zeta + \frac{1}{2\zeta} = 1,414 \tag{7.15}$$

Se $c_2^2 > 0,62c_1^2$, então a Eq. (7.14) se torna

$$\zeta = \sqrt{\frac{c_1^2 + c_2^2}{2c_1^2}} > \sqrt{\frac{1}{2} + \frac{0,62}{2}} = 0,9$$

Mas ζ está limitado entre 0 e 0,9. Portanto a Eq. (7.14) não fornecerá o valor ótimo de ζ. Para se determinar o valor ótimo de ζ, definimos

$$c_2^2 = (0,62 + 2h)c_1^2, \qquad h > 0$$

Então podemos escrever a Eq. (7.13) como segue:

$$J = \left(\zeta + \frac{0,81 + h}{\zeta}\right)c_1^2 + c_1c_2$$

O valor ótimo de ζ que minimiza J para dados valores de c_1 e c_2 pode ser obtido minimizando-se a seguinte expressão I com relação a ζ.

$$I = \zeta + \frac{0,81 + h}{\zeta}$$

O valor de I é mínimo quando $\zeta^2 = 0,81 + h$. Entretanto, este valor de ζ não pode ser obtido por causa da condição $0 < \zeta \leq 0,9$. Como o valor de I aumenta quando ζ diminui a partir de $\sqrt{0,81 + h}$, o valor ótimo de ζ quando $c_2^2 > 0,62c_1^2$ é 0,9.

Resumindo, o valor ótimo de ζ é

$$\zeta = \sqrt{\frac{c_1^2 + c_2^2}{2c_1^2}} \qquad \text{se } c_2^2 \leq 0,62c_1^2$$
$$= 0,9 \qquad \text{se } c_2^2 > 0,62c_1^2$$

Problema A.7.6 Calcule os seguintes índices de desempenho:

$$\int_0^\infty |e(t)|\,dt,$$

$$\int_0^\infty t\,|e(t)|\,dt,$$

$$\int_0^\infty t[|e(t)| + |\dot{e}(t)|]\,dt$$

para o sistema

$$\frac{C(s)}{R(s)} = \frac{1}{s^2 + 2\zeta s + 1} \qquad (\zeta \geq 1)$$
$$E(s) = \mathscr{L}[e(t)] = R(s) - C(s)$$

Suponha que o sistema está inicialmente em repouso e é sujeito a uma entrada degrau unitário.

Solução. Para uma entrada degrau unitário, $R(s) = 1/s$ e

$$E(s) = \frac{s^2 + 2\zeta s}{s^2 + 2\zeta s + 1}\frac{1}{s}$$
$$= \frac{-\zeta + \sqrt{\zeta^2 - 1}}{2\sqrt{\zeta^2 - 1}}\frac{1}{s + \zeta + \sqrt{\zeta^2 - 1}} + \frac{\zeta + \sqrt{\zeta^2 - 1}}{2\sqrt{\zeta^2 - 1}}\frac{1}{s + \zeta - \sqrt{\zeta^2 - 1}}$$

Para $\zeta \geq 1$, o sistema não apresenta sobre-sinal. Portanto $|e(t)| = e(t)$ para todo $t \geq 0$. Os dados índices de desempenho podem então ser computados como segue:

1. $$\int_0^\infty |e(t)|\,dt = \int_0^\infty e(t)\,dt$$
$$= \lim_{s\to 0} E(s)$$
$$= \frac{-\zeta + \sqrt{\zeta^2 - 1}}{2\sqrt{\zeta^2 - 1}}\frac{1}{\zeta + \sqrt{\zeta^2 - 1}} + \frac{\zeta + \sqrt{\zeta^2 - 1}}{2\sqrt{\zeta^2 - 1}}\frac{1}{\zeta - \sqrt{\zeta^2 - 1}}$$
$$= 2\zeta \qquad (\zeta \geq 1)$$

2. $$\int_0^\infty t\,|e(t)|\,dt = \int_0^\infty te(t)\,dt$$
$$= \lim_{s\to 0}\left[-\frac{d}{ds}E(s)\right]$$
$$= \lim_{s\to 0}\left[\frac{-\zeta + \sqrt{\zeta^2 - 1}}{2\sqrt{\zeta^2 - 1}}\frac{1}{(s + \zeta + \sqrt{\zeta^2 - 1})^2} + \frac{\zeta + \sqrt{\zeta^2 - 1}}{2\sqrt{\zeta^2 - 1}}\frac{1}{(s + \zeta - \sqrt{\zeta^2 - 1})^2}\right]$$
$$= 4\zeta^2 - 1 \qquad (\zeta \geq 1)$$

3. Observando que $\dot{e}(t) < 0$ e $|\dot{e}(t)| = -\dot{e}(t)$ para $t \geq 0$, obtemos

$$\int_0^\infty t(|e(t)| + |\dot{e}(t)|)\,dt = \int_0^\infty [te(t) - t\dot{e}(t)]\,dt$$
$$= \lim_{s\to 0}\left[-\frac{d}{ds}E(s) + \frac{d}{ds}sE(s)\right]$$
$$= \lim_{s\to 0}\left[(-1 + s)\frac{d}{ds}E(s) + E(s)\right]$$
$$= 4\zeta^2 - 1 + 2\zeta \qquad (\zeta \geq 1)$$

Para fins de comparação, forneceremos alguns outros índices de desempenho calculados para o mesmo sistema. [Referir-se às Eqs. (7.9), (7.10) e (7.15).]

$$\int_0^\infty e^2(t)\,dt = \zeta + \frac{1}{4\zeta} \qquad (\zeta > 0)$$
$$\int_0^\infty te^2(t)\,dt = \zeta^2 + \frac{1}{8\zeta^2} \qquad (\zeta > 0)$$
$$\int_0^\infty [e^2(t) + \dot{e}^2(t)]\,dt = \zeta + \frac{1}{2\zeta} \qquad (\zeta > 0)$$

Gráficos destes seis índices de desempenho são vistos na Fig. 7.8 em função de ζ. (Note que as curvas vistas na Fig. 7.8 são válidas para $\zeta > 0$. Os três índices de desempenho calculados neste problema são válidos apenas para $\zeta \geq 1$.)

PROBLEMAS

Problema B.7.1 Mostre que o erro estacionário na resposta a entradas em rampa pode ser feito nulo se a função de transferência de malha-fechada é dada por

$$\frac{C(s)}{R(s)} = \frac{a_{n-1}s + a_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}$$

Problema B.7.2 Obtenha os coeficientes de erro de velocidade dinâmico dos seguintes sistemas:

1. $$\frac{C(s)}{R(s)} = \frac{10}{(s + 1)(5s^2 + 2s + 10)}$$

2. $$\frac{C(s)}{R(s)} = \frac{3s + 10}{5s^2 + 2s + 10}$$

Problema B.7.3 Considere o sistema de controle com realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{100}{s(0{,}1s + 1)}$$

Determine o erro estacionário quando a entrada é

$$r(t) = 1 + t + at^2, \qquad a \geq 0$$

Problema B.7.4 Considere o sistema visto na Fig. 7.11. Suponha que a entrada é uma entrada em rampa, ou

$$r(t) = at$$

onde a é uma constante arbitrária.

Mostre que ajustando-se adequadamente o valor de K_i, o erro estacionário na resposta a entradas em rampa pode ser feito nulo.

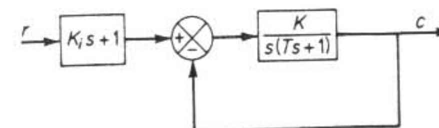


Fig. 7.11 Sistema de controle.

Problema B.7.5 Considere um sistema descrito por

$$\ddot{x} + 2\zeta\dot{x} + x = 0$$

Determine o valor do coeficiente de amortecimento ζ de tal forma que

$$\int_0^\infty (x^2 + \mu \dot{x}^2)\, dt \qquad (\mu \geq 0)$$

seja minimizado.

Problema B.7.6 A função de transferência de malha-fechada para uma entrada de perturbação $n(t)$ é dada por

$$\frac{C(s)}{N(s)} = \frac{s(s + a)}{s^2 + as + 10}$$

Determine o valor da constante a de tal forma que a integral do erro quadrático devido a uma perturbação em degrau seja minimizada.

***Problema B.7.7** Considere um sistema de controle com realimentação unitária com a função de transferência de malha-aberta

$$G(s) = \frac{(a + b)s + ab}{s^2}$$

Para uma entrada degrau unitário, calcule os seguintes índices de desempenho:

$$J_1 = \int_0^\infty e^2(t)\, dt$$

$$J_2 = \int_0^\infty te^2(t)\, dt$$

Se $a = b$, quais são os valores de J_1 e J_2?

8

O Método do Lugar das Raízes

8.1 INTRODUÇÃO

A característica básica da resposta transitória de um sistema em malha-fechada é determinada a partir dos pólos da malha-fechada. Portanto, em problemas de análise, é importante localizar os pólos em malha-fechada no plano s. No projeto de sistemas em malha-fechada, queremos ajustar os pólos e zeros de malha-aberta de modo a colocar os pólos e zeros em malha-fechada nas posições desejadas do plano s.

Os pólos de malha-fechada são as raízes da equação característica. Para determiná-los necessitamos fatorar o polinômio característico. Em geral, este procedimento é trabalhoso se o grau do polinômio característico é três, ou maior. As técnicas clássicas de fatoração de polinômios não são convenientes porque conforme varia o ganho da função de transferência em malha-aberta, devem ser repetidos os cálculos.

Um método simples para determinar as raízes da equação característica foi desenvolvido por W. R. Evans e é extensivamente usado em engenharia de controle. Este método, denominado *método do lugar das raízes*, é um método pelo qual as raízes da equação característica são colocadas em um gráfico para todos os valores de um parâmetro do sistema. As raízes correspondentes a um valor particular deste parâmetro podem então ser localizadas no gráfico resultante. Note que o parâmetro usualmente é o ganho, porém qualquer outra variável da função de transferência em malha-aberta pode ser utilizada. Salvo menção em contrário, suporemos que o ganho da função de transferência em malha-fechada é o parâmetro a ser variado através de todos os seus valores, isto é, de zero a infinito.

Método do lugar das raízes. A idéia básica na qual se baseia o método do lugar das raízes é que o valor de s que faz a função de transferência pelo laço ser igual a -1 deve satisfazer a equação característica do sistema.

O lugar das raízes da equação característica do sistema em malha-fechada, conforme o ganho é variado desde zero até infinito, dá ao método o seu nome. O

gráfico correspondente mostra claramente as contribuições de cada pólo ou zero de malha-aberta nas localizações dos pólos em malha-fechada.

O método do lugar das raízes nos possibilita determinar os pólos em malha-fechada a partir dos pólos e zeros de malha-aberta, considerando o ganho como parâmetro. Conseqüentemente, evita dificuldades inerentes às técnicas clássicas, fornecendo uma imagem gráfica de todos os pólos de malha-fechada para todos os valores do ganho da função de transferência em malha-aberta.

No projeto de um sistema de controle linear, verificamos que o método do lugar das raízes se torna muito útil desde que indica a maneira pela qual os pólos e zeros em malha-aberta devem ser modificados de modo que a resposta satisfaça as especificações de desempenho do sistema. Este método é particularmente conveniente para se obter resultados aproximados muito rapidamente.

Desde que o método é gráfico na determinação das raízes da equação característica, ele fornece um procedimento gráfico eficaz para determinar as raízes de qualquer equação polinomial que resulte de um estudo de sistemas físicos.

Esboço do capítulo. Na Seção 8.2 introduziremos os conceitos que sustentam o método do lugar das raízes. A Seção 8.3 apresenta o procedimento geral para esboçar os lugares das raízes. Seguindo esta apresentação, na Seção 8.4 analisaremos sistemas de malha-fechada pelo uso do método do lugar das raízes. (A abordagem de lugar das raízes no projeto de sistemas será discutida no Cap. 10.)

8.2 DIAGRAMAS DE LUGAR DAS RAÍZES

Condições de ângulo e amplitude. Considere o sistema indicado na Fig. 8.1. A função de transferência em malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} \tag{8.1}$$

A equação característica para este sistema em malha-fechada é obtida igualando-se o denominador da fração do segundo membro da Eq. (8.1) a zero. Isto é,

$$1 + G(s)H(s) = 0$$

ou

$$G(s)H(s) = -1 \tag{8.2}$$

Desde que $G(s)H(s)$ é uma quantidade complexa, a Eq. (8.2) deve ser desmembrada em duas equações a fim de se igualar os ângulos e os módulos de ambos os membros da equação, respectivamente, para obter

Condição de ângulo:

$$\angle G(s)H(s) = \pm 180°(2k + 1) \qquad (k = 0, 1, 2, \ldots) \tag{8.3}$$

Condição de módulo:

$$|G(s)H(s)| = 1 \tag{8.4}$$

Os valores de s que satisfazem as condições de ângulo e módulo são as raízes da equação característica, ou os pólos de malha-fechada. Um gráfico (ou diagrama) dos pontos do plano complexo que satisfazem apenas a condição do ângulo é o lugar

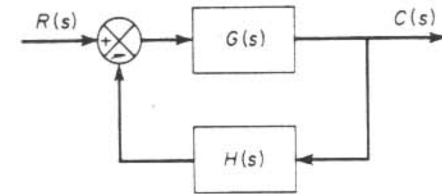


Fig. 8.1 Sistema de controle.

das raízes. As raízes na equação característica (os pólos em malha-fechada) correspondentes a um dado valor do ganho podem ser determinadas a partir da condição de módulo. Os detalhes da aplicação das condições de ângulo e módulo a fim de se obter os pólos em malha-fechada estão apresentados na Seção 8.3.

Gráficos de lugar das raízes para sistemas de segunda-ordem. Antes de apresentarmos um método para construção destes gráficos em detalhes, será ilustrado um gráfico do lugar das raízes para um sistema simples de segunda-ordem.

Considere o sistema indicado na Fig. 8.2. A função de transferência em malha-aberta $G(s)H(s)$ é

$$G(s)H(s) = \frac{K}{s(s+1)}$$

A função de transferência em malha-fechada é então

$$\frac{C(s)}{R(s)} = \frac{K}{s^2 + s + K}$$

A equação característica é

$$s^2 + s + K = 0 \tag{8.5}$$

Desejamos determinar o lugar das raízes desta equação conforme K varia desde zero até infinito.

A fim de ilustrar claramente qual o aspecto do gráfico do lugar das raízes para este sistema, obteremos inicialmente as raízes da equação característica analiticamente em termos de K, e então variaremos K desde zero até infinito. Deve ser notado que este não é o modo apropriado para construir o gráfico do lugar das raízes. O modo apropriado é através de uma abordagem gráfica de tentativa-e-erro, e o trabalho gráfico pode ser muito simplificado aplicando-se as regras gerais a serem apresentadas na Seção 8.3. (Obviamente, se puder ser encontrada facilmente uma solução analítica para as raízes da equação característica, não há necessidade

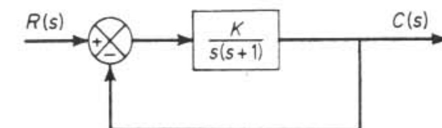


Fig. 8.2 Sistema de controle.

de se utilizar o método do lugar das raízes.) As raízes da equação característica, Eq. (8.5), são

$$s_1 = -\tfrac{1}{2} + \tfrac{1}{2}\sqrt{1-4K}, \qquad s_2 = -\tfrac{1}{2} - \tfrac{1}{2}\sqrt{1-4K}$$

As raízes são reais para $K \leq 1/4$ e complexas para $K > 1/4$.

O lugar das raízes correspondente a todos os valores de K está indicado na Fig. 8.3(a). O lugar das raízes está graduado com K como parâmetro. (O movimento das raízes conforme K aumenta é indicado por setas.) Uma vez desenhado o gráfico, podemos imediatamente determinar o valor de K que fornecerá uma raiz, ou um pólo em malha-fechada, em um ponto desejado. Desta análise, é claro que os pólos em malha-fechada correspondentes a $K = 0$ são os mesmos pólos de $G(s)H(s)$. Conforme o valor de K aumenta desde zero até 1/4, os pólos de malha-fechada movem-se para o ponto $(-1/2; 0)$. Para valores de K entre zero e 1/4, todos os pólos de malha-fechada estão sobre o eixo real. Esta é a situação correspondente a um sistema superamortecido, e a resposta impulsiva não é oscilatória. Em $K = 1/4$, os dois pólos de malha-fechada reais se igualam. Esta situação corresponde ao caso do sistema criticamente amortecido. Conforme K aumenta a partir de 1/4, os pólos de malha-fechada tornam-se complexos, posicionando-se fora do eixo real, e devido à parte real dos pólos de malha-fechada ser constante para $K > 1/4$, os pólos de malha-fechada movem-se ao longo da reta $s = -1/2$. Portanto, para $K > 1/4$, o sistema comporta-se como subamortecido. Para um dado valor de K, um dos pólos conjugados de malha-fechada move-se para $s = -1/2 + j\infty$, enquanto o outro move-se para $s = -1/2 - j\infty$.

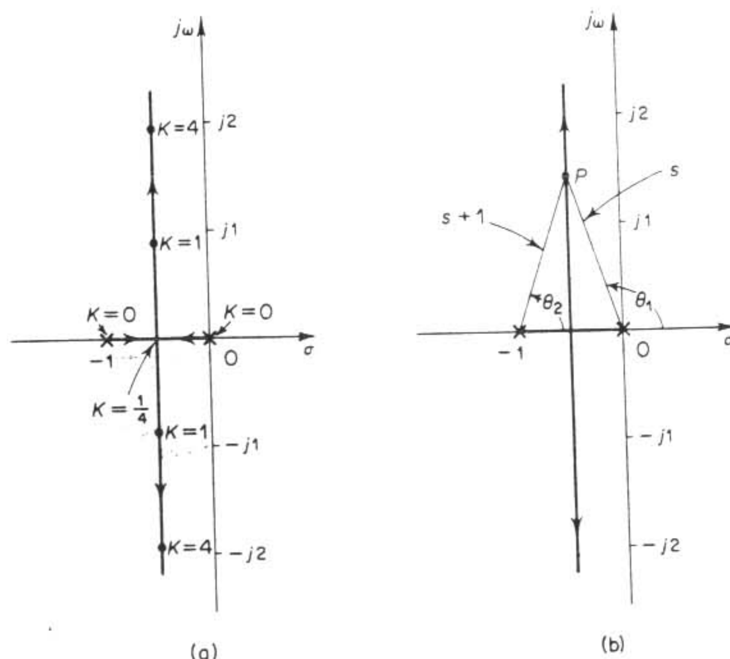


Fig. 8.3 Gráficos do lugar das raízes dos sistemas indicados na Fig. 8.2.

Mostraremos que qualquer ponto sobre o lugar das raízes satisfaz a condição de ângulo. A condição de ângulo dada pela Eq. (8.3) é

$$\angle\frac{K}{s(s+1)} = -\angle s - \angle s+1 = \pm 180°(2k+1) \qquad (k = 0, 1, 2, \ldots)$$

Considere o ponto P sobre o lugar das raízes indicado na Fig. 8.3(b). Os números complexos s e $s + 1$ possuem ângulos θ_1 e θ_2, respectivamente, e módulos $|s|$ e $|s + 1|$, respectivamente. (Note que todos os ângulos são considerados positivos quando são medidos no sentido anti-horário.) A soma dos ângulos θ_1 e θ_2 nitidamente é 180°.

Se o ponto P estiver localizado sobre o eixo real entre 0 e −1, então $\theta_1 = 180°$ e $\theta_2 = 0°$. Portanto, verificamos que qualquer ponto sobre o lugar das raízes satisfaz a condição de ângulo. Notemos também que se o ponto P não for um ponto sobre o lugar das raízes, então a soma de θ_1 e θ_2 não é igual a $\pm 180°(2k + 1)$, onde $k = 0, 1, 2, \ldots$ Portanto, os pontos que não estiverem sobre o lugar das raízes não satisfazem a condição de ângulo e, portanto, não podem ser pólos de malha-fechada para quaisquer valores de K.

Se os pólos de malha-fechada forem especificados no lugar das raízes, então o valor correspondente de K é determinado pela condição do módulo, Eq. (8.4). Se, por exemplo, os pólos de malha-fechada selecionados são $s = -1/2 \pm j2$, então o valor correspondente de K é determinado de

$$|G(s)H(s)| = \left|\frac{K}{s(s+1)}\right|_{s=-(1/2)+j2} = 1$$

ou

$$K = |s(s+1)|_{s=-(1/2)+j2} = \frac{17}{4}$$

Desde que pólos complexos são conjugados, se um deles, por exemplo, $s = -1/2 + j2$, é especificado, então o outro é automaticamente fixado. No cálculo do valor de K, qualquer um dos pólos pode ser utilizado.

Do gráfico do lugar das raízes fornecido na Fig. 8.3(a), verificamos claramente os efeitos de variações no valor de K no comportamento da resposta transitória em relação ao sistema de segunda-ordem. Um aumento no valor de K diminuirá a relação de amortecimento ζ, resultando em um aumento na sobrelevação da resposta. Um aumento no valor de K também resultará em um aumento nas freqüências natural não amortecida e amortecida. (Se K é maior do que o valor crítico, que corresponde a um sistema criticamente amortecido, um aumento no valor de K não afetará o valor da parte real dos pólos em malha-fechada.) A partir do gráfico do lugar das raízes, é evidente que os pólos de malha-fechada sempre estarão no semiplano esquerdo do plano s; conseqüentemente, independentemente de quanto o parâmetro K for aumentado, o sistema sempre permanecerá estável; desta forma, o sistema de segunda-ordem é sempre estável. (Note, entretanto, que se o ganho for ajustado em um valor muito alto, os efeitos de alguma das constantes de tempo que foram desprezadas podem tornar-se importantes, e o sistema que é supostamente de segunda-ordem, porém efetivamente de ordem maior, pode tornar-se instável.)

A Tabela 8.1 fornece uma coleção de gráficos simples de lugares de raízes.

Tabela 8.1 Coleção de gráficos de lugares das raízes simples

$G(s)H(s)$	Localizações do pólo-zero de malha-aberta e lugares das raízes	$G(s)H(s)$	Localizações do pólo-zero de malha-aberta e lugares das raízes
$\frac{K}{s}$	$j\omega$; σ	$\frac{K}{s^2}$	$j\omega$; σ
$\frac{K}{s+p}$	$j\omega$; $-p$; σ	$\frac{K}{s^2+\omega_1^2}$	$j\omega$; $j\omega_1$; σ; $-j\omega_1$
$\frac{K(s+z)}{s+p}$ $(z > p)$	$j\omega$; $-z$ $-p$; σ	$\frac{K}{(s+\sigma)^2+\omega_1^2}$	$j\omega$; $j\omega_1$; $-\sigma$; σ; $-j\omega_1$
$\frac{K(s+z)}{s+p}$ $(z < p)$	$j\omega$; $-p$ $-z$; σ	$\frac{K}{(s+p_1)(s+p_2)}$	$j\omega$; $-p_1$; $-p_2$; σ

Lugares de ganho constante. A Fig. 8.4 mostra um gráfico dos lugares de ganho constante do sistema indicado na Fig. 8.2.

Os lugares de ganho constante para este sistema foram obtidos a partir da condição de módulo:

$$|G(s)H(s)| = \left|\frac{K}{s(s+1)}\right| = 1$$

ou

$$|s(s+1)| = K \tag{8.6}$$

Os pontos do plano complexo que satisfazem a Eq. (8.6) para um dado K constituem um lugar de ganho constante.

Ortogonalidade entre os lugares das raízes e os lugares de ganho constante. Considere o sistema indicado na Fig. 8.1. No plano $G(s)H(s)$ os lugares de $|G(s)H(s)| =$ constante são circunferências com centro na origem, e os lugares correspondentes a $\underline{/G(s)H(s)} = \pm 180°(2k+1)$ $(k = 0, 1, 2,...)$ permanecem sobre o eixo real negativo do

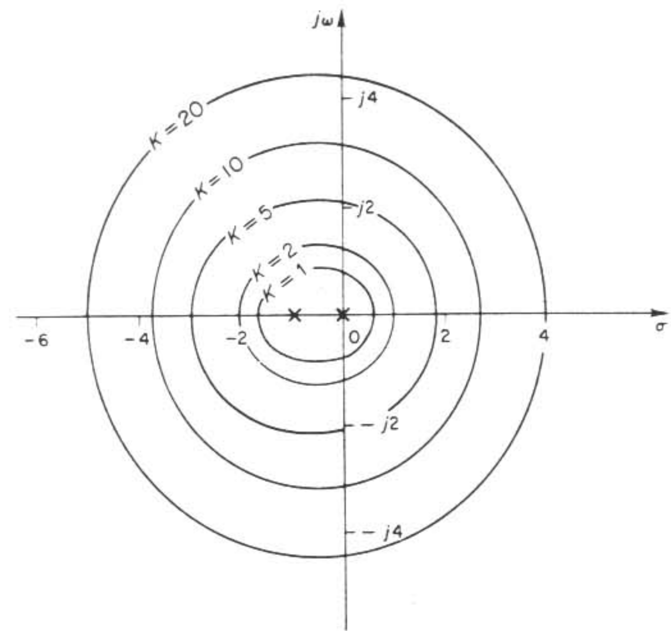


Fig. 8.4 Gráfico dos lugares de ganho constante do sistema indicado na Fig. 8.2.

plano $G(s)H(s)$, como indicado na Fig. 8.5. [Note que o plano complexo aqui empregado não é o plano s, mas o plano $G(s)H(s)$.]

Os lugares das raízes e os lugares de ganho constante no plano s constituem mapeamentos conformes dos lugares de $\underline{/G(s)H(s)} = \pm 180°(2k+1)$ e de $|G(s)H(s)| =$ constante no plano $G(s)H(s)$.

Desde que os lugares de fase constante e ganho constante no plano $G(s)H(s)$ são ortogonais, os lugares das raízes e os lugares de ganho constante no plano s também são ortogonais. A Fig. 8.6(a) mostra os lugares das raízes e os lugares de ganho constante para o seguinte sistema:

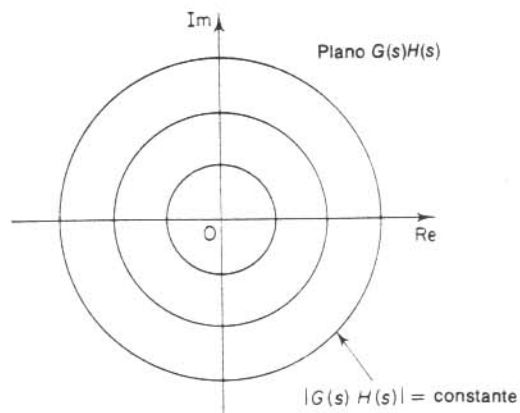
$$G(s) = \frac{K(s+2)}{s^2+2s+3}, \qquad H(s) = 1$$

Verifique que devido à configuração de pólo-zero ser simétrica em relação ao eixo real, os lugares de ganho constante também são simétricos em relação ao eixo real.

A Fig. 8.6(b) mostra os lugares das raízes e de ganho constante para o sistema:

$$G(s) = \frac{K}{s(s+1)(s+2)}, \qquad H(s) = 1$$

Note que devido à configuração dos pólos no plano s ser simétrica em relação ao eixo real, e em relação à reta paralela ao eixo imaginário passando pelo ponto ($\sigma = -1, \omega = 0$), os lugares de ganho constante são simétricos em relação à reta $\omega = 0$ (eixo real) e à reta $\sigma = -1$.



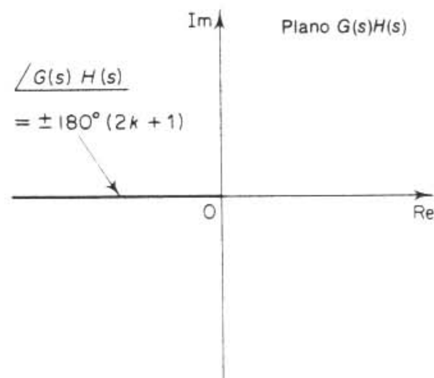
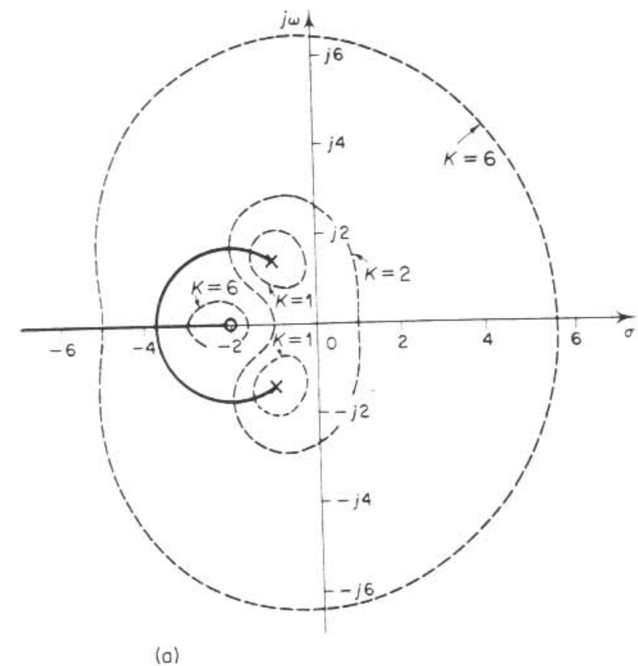


Fig. 8.5 Gráficos dos lugares de ganho constante e fase constante no plano $G(s)H(s)$.



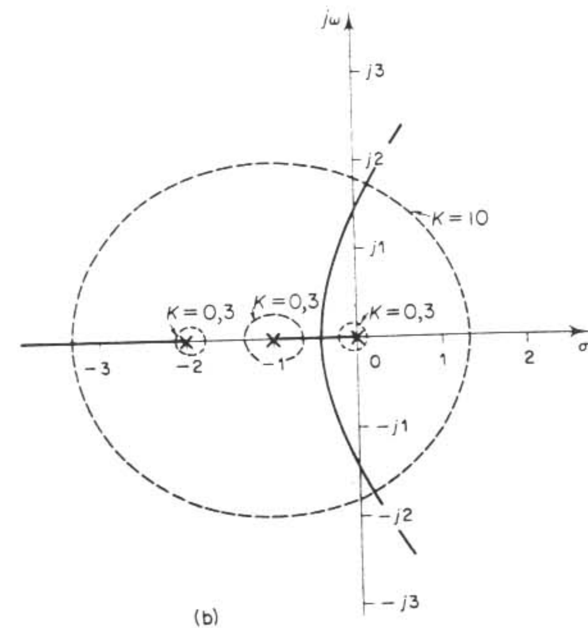


Fig. 8.6 Gráficos dos lugares das raízes e dos lugares de ganho constante.

8.3 DOIS EXEMPLOS ILUSTRATIVOS

Nesta seção apresentaremos dois exemplos ilustrativos para construção dos gráficos de lugar das raízes. Utilizaremos determinações gráficas, combinadas com inspeção, para determinar os lugares das raízes sobre os quais as raízes da equação característica do sistema de malha-fechada devem permanecer. Embora empreguemos apenas sistemas simples para fins de ilustração, o procedimento para determinar os lugares das raízes para sistemas de ordem maior não é mais complicado.

O primeiro passo no procedimento para construção do gráfico do lugar das raízes é determinar os lugares das possíveis raízes utilizando a condição de ângulo. Posteriormente, os lugares das raízes são calibrados ou graduados usando-se a condição de módulo.

Como as medidas gráficas de ângulos são envolvidas na análise, esboçando-se o lugar das raízes em um papel de gráfico, verifica-se a necessidade de usar as mesmas escalas nos eixos das abscissas e das ordenadas.

Exemplo 8.1 Considere o sistema indicado na Fig. 8.7. (Suporemos que o valor do ganho K é positivo.) Para este sistema,

$$G(s) = \frac{K}{s(s+1)(s+2)}, \qquad H(s) = 1$$

Esbocemos o gráfico do lugar das raízes e determinemos posteriormente o valor de K para que a relação de amortecimento ζ de um par de pólos de malha-fechada complexos conjugados dominantes seja 0.5.

Para o sistema dado, a condição de ângulo fornece

$$\begin{aligned}\angle G(s) &= \angle \frac{K}{s(s+1)(s+2)} \\ &= -\angle s - \angle s+1 - \angle s+2 \\ &= \pm 180^\circ(2k+1) \qquad (k = 0, 1, 2, \ldots)\end{aligned}$$

A condição de módulo é

$$|G(s)| = \left|\frac{K}{s(s+1)(s+2)}\right| = 1$$

Um procedimento típico para esboçar o gráfico do lugar das raízes é o seguinte:

1. Determine os lugares das raízes no eixo real. O primeiro passo na construção de um gráfico do lugar das raízes é localizar os pólos de malha-aberta, $s = 0$, $s = -1$, e $s = -2$ no plano complexo. (Não há zeros de malha-aberta neste sistema.) As posições dos pólos de

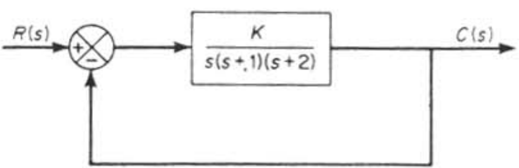


Fig. 8.7 Sistema de controle.

malha-aberta são indicadas por cruzes. (As posições dos zeros de malha-aberta serão indicadas neste livro por pequenas circunferências.) Note que os pontos de partida dos lugares das raízes (os pontos correspondentes a $K = 0$), são os pólos de malha-aberta. O número de lugares das raízes individuais para este sistema é três, que é igual ao número de pólos de malha-aberta.

Para determinar os lugares das raízes sobre o eixo real, selecionamos um ponto de teste s. Se o ponto de teste estiver sobre o eixo real positivo, então

$$\angle s = \angle s+1 = \angle s+2 = 0^\circ$$

Isto indica que a condição de ângulo não pode ser satisfeita. Conseqüentemente, não há nenhum lugar das raízes sobre o eixo real positivo. Posteriormente, selecione um ponto de teste sobre o eixo real negativo entre 0 e −1. Então,

$$\angle s = 180^\circ, \qquad \angle s+1 = \angle s+2 = 0^\circ$$

Portanto,

$$-\angle s - \angle s+1 - \angle s+2 = -180^\circ$$

e a condição do ângulo é satisfeita. Portanto, a parte do eixo real negativo entre 0 e −1 constitui uma parte do lugar das raízes. Se um ponto de teste for selecionado entre −1 e −2, então,

$$\angle s = \angle s+1 = 180^\circ, \qquad \angle s+2 = 0^\circ$$

Verifica-se então que a condição de ângulo não será satisfeita. Portanto, o eixo real negativo desde −1 até −2 não corresponde a uma parte do lugar das raízes. Analogamente, se um ponto de teste for localizado no eixo real negativo desde −2 até $-\infty$, a condição de ângulo será satisfeita. Conseqüentemente, existirá um lugar das raízes sobre o eixo real negativo entre 0 e −1 e entre −2 e $-\infty$.

2. Determine as assíntotas dos lugares das raízes. As assíntotas dos lugares das raízes, conforme s tende a infinito, podem ser determinadas como segue: Se um ponto de teste é selecionado muito distante da origem, então

$$\lim_{s\to\infty} G(s) = \lim_{s\to\infty} \frac{K}{s(s+1)(s+2)} = \lim_{s\to\infty} \frac{K}{s^3}$$

e a condição de ângulo fornece

$$-3\angle s = \pm 180^\circ(2k+1) \qquad (k = 0, 1, 2, \ldots)$$

ou

$$\text{Ângulo da assíntota} = \frac{\pm 180^\circ(2k+1)}{3}$$

Desde que o ângulo se repete conforme k é variado, os ângulos distintos para as assíntotas são apenas 60°, −60° e 180°. Portanto, há três assíntotas. Aquela correspondente ao ângulo de 180° está sobre o eixo real negativo.

Antes de desenhar estas assíntotas no plano complexo, devemos determinar onde elas interceptam o eixo real. Desde que a equação característica deste sistema é

$$\frac{K}{s(s+1)(s+2)} = -1$$

ou

$$s^3 + 3s^2 + 2s = -K$$

se s é suposto muito grande, podemos então aproximar esta equação característica pela seguinte equação:

$$(s + 1)^3 = 0$$

A abscissa da intersecção da assíntota no eixo real pode ser obtida substituindo-se $s = \sigma_a$ e determinando-se a solução da equação para σ_a. Portanto,

$$\sigma_a = -1$$

e o ponto da origem das assíntotas é $(-1, 0)$. Estas assíntotas constituem partes do lugar das raízes nas regiões muito distantes da origem.

3. Determine o ponto de separação de partida. Para desenhar precisamente o gráfico do lugar das raízes devemos determinar o ponto de separação de partida, onde os ramos do lugar das raízes originários dos pólos 0 e -1 separam-se (conforme K é aumentado) do eixo real e movem-se para o plano complexo em pontos com parte imaginária não nula. O ponto de separação corresponde a um ponto no plano s onde são encontradas raízes múltiplas da equação característica.

O ponto de separação $(-\sigma_b, 0)$ pode ser determinado como segue: Suponha um ponto de teste que esteja próximo ao eixo real negativo entre 0 e -1 conforme indicado na Fig. 8.8. Se denotarmos a distância vertical do ponto de teste, medida a partir do eixo real negativo, por δ, então os ângulos θ_1, θ_2 e θ_3 dos números complexos s, $s + 1$ e $s + 2$, respectivamente, podem ser escritos como segue:

$$\theta_1 = 180° - \tan^{-1}\left(\frac{\delta}{\sigma_b}\right)$$

$$\theta_2 = \tan^{-1}\left[\frac{\delta}{-\sigma_b - (-1)}\right]$$

$$\theta_3 = \tan^{-1}\left[\frac{\delta}{-\sigma_b - (-2)}\right]$$

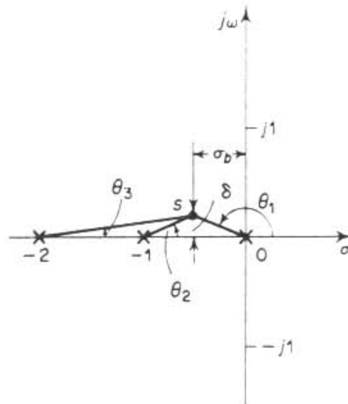


Fig. 8.8 Determinação do ponto de separação de partida.

Para um valor pequeno de δ,

$$\theta_1 = 180° - \frac{\delta}{\sigma_b}$$

$$\theta_2 = \frac{\delta}{-\sigma_b + 1}$$

$$\theta_3 = \frac{\delta}{-\sigma_b + 2}$$

A condição de ângulo

$$\theta_1 + \theta_2 + \theta_3 = 180°$$

resulta

$$-\frac{\delta}{\sigma_b} + \frac{\delta}{-\sigma_b + 1} + \frac{\delta}{-\sigma_b + 2} = 0°$$

ou

$$\frac{1}{\sigma_b} + \frac{1}{\sigma_b - 1} + \frac{1}{\sigma_b - 2} = 0 \quad (8.7)$$

Resolvendo a Eq. (8.7), obtemos

$$3\sigma_b^2 - 6\sigma_b + 2 = 0$$

resultando

$$\sigma_b = 0{,}423, \qquad \sigma_b = 1{,}577$$

Desde que $0 > -\sigma_b > -1$, o ponto de separação de partida deve ser $(-0{,}423, 0)$. Se o número de termos da Eq. (8.7) for quatro ou mais, deve ser utilizada uma abordagem de tentativa-e-erro para determinar o valor de σ_b. Esta abordagem pode ser facilitada se determinarmos um valor aproximado para σ_b desprezando os pólos e zeros localizados distantes do ponto de separação de partida, e posteriormente fazendo uma correção para este valor aproximado de σ_b.

É disponível um método alternativo para determinação do ponto de separação de partida. Apresentaremos este método a seguir: Vamos escrever a equação característica como

$$f(s) = A(s) + KB(s) = 0 \quad (8.8)$$

onde $A(s)$ e $B(s)$ não contêm K. Note que $f(s) = 0$ possui raízes múltiplas nos pontos onde

$$\frac{df(s)}{ds} = 0$$

Da Eq. (8.8), obtemos

$$\frac{df(s)}{ds} = A'(s) + KB'(s) = 0 \quad (8.9)$$

onde

$$A'(s) = \frac{dA(s)}{ds},$$

$$B'(s) = \frac{dB(s)}{ds}$$

O valor particular de K que fornece raízes múltiplas da equação característica é obtido da Eq. (8.9) como

$$K = -\frac{A'(s)}{B'(s)}$$

Se substituirmos este valor de K na Eq. (8.8), obtemos

$$f(s) = A(s) - \frac{A'(s)}{B'(s)}B(s) = 0$$

ou

$$A(s)B'(s) - A'(s)B(s) = 0 \tag{8.10}$$

Se a Eq. (8.10) for resolvida em relação a s, os pontos onde ocorrem raízes múltiplas podem ser obtidos. Por outro lado, da Eq. (8.8) obtemos

$$K = -\frac{A(s)}{B(s)}$$

e

$$\frac{dK}{ds} = -\frac{A'(s)B(s) - A(s)B'(s)}{B^2(s)}$$

Se dK/ds é igualado a zero, obtemos uma equação idêntica à Eq. (8.10). Portanto, os pontos de separação de partida podem ser simplesmente determinados das raízes de

$$\frac{dK}{ds} = 0$$

Deve ser notado que nem todas as soluções da Eq. (8.10) ou de $dK/ds = 0$ correspondem a pontos de separação de partida verdadeiros. Demonstraremos este fato utilizando o exemplo presente, onde

$$f(s) = s^3 + 3s^2 + 2s + K = 0 \tag{8.11}$$

Referindo-se ao gráfico de $f(\sigma)$ *versus* σ indicado na Fig. 8.9, note que os pontos P e Q correspondem a $df(\sigma)/d\sigma = 0$. É claro que o ponto Q corresponde a $K < 0$, o que significa que o ponto Q não pode ser um ponto de separação de partida do sistema sob consideração. (Se num ponto no qual $df(s)/ds = 0$ o valor de K corresponder a um valor real positivo, então o ponto considerado é um ponto de separação de partida verdadeiro.)

Para o presente exemplo, da Eq. (8.11) obtemos

$$K = -s(s + 1)(s + 2) = -(s^3 + 3s^2 + 2s)$$

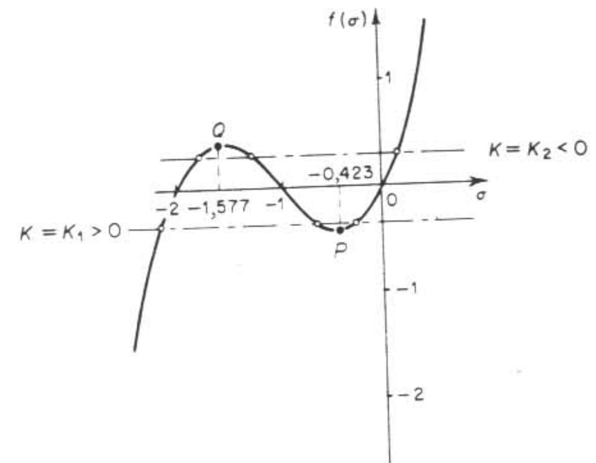


Fig. 8.9 Gráfico de $f(\sigma)$ versus σ.

Impondo que $dK/ds = 0$, obtemos

$$\frac{dK}{ds} = -(3s^2 + 6s + 2) = 0$$

ou

$$s = -0{,}423, \qquad s = -1{,}577$$

Desde que o ponto de separação de partida deve estar entre 0 e -1, é claro que $s = -0{,}423$ corresponde a um ponto de separação de partida verdadeiro. De fato, o cálculo dos valores de K correspondentes a $s = -0{,}423$ e $s = -1{,}577$ fornece

$$K = 0{,}385 \qquad \text{para } s = -0{,}423$$
$$K = -0{,}385 \qquad \text{para } s = -1{,}577$$

4. Determine os pontos onde os lugares das raízes cruzam o eixo imaginário. Estes pontos podem ser determinados utilizando-se o critério de estabilidade de Routh como segue. Desde que a equação característica para o presente sistema é

$$s^3 + 3s^2 + 2s + K = 0$$

o arranjo de Routh resulta em

$$\begin{array}{lll} s^3 & 1 & 2 \\ s^2 & 3 & K \\ s^1 & \frac{6 - K}{3} & \\ s^0 & K & \end{array}$$

O valor de K que faz com que o termo s^1 na primeira coluna seja igual a zero é $K = 6$. Os pontos do cruzamento do eixo imaginário podem ser determinados pela solução da equação auxiliar obtida a partir da linha s^2; isto é,

$$3s^2 - K = 3s^2 + 6 = 0$$

que fornece

$$s = \pm j\sqrt{2}$$

As freqüências dos pontos de cruzamento no eixo imaginário são portanto $\omega = \pm\sqrt{2}$. O valor do ganho correspondente aos pontos de cruzamento é $K = 6$.

Uma abordagem alternativa é substituir $s = j\omega$ na equação característica, igualar tanto a parte real como a parte imaginária a zero, e então resolver para ω e K. Para o sistema dado, a equação característica, com $s = j\omega$, é

$$(j\omega)^3 + 3(j\omega)^2 + 2(j\omega) + K = 0$$

ou

$$(K - 3\omega^2) + j(2\omega - \omega^3) = 0$$

Igualando-se a zero as partes real e imaginária desta última equação, obtemos

$$K - 3\omega^2 = 0, \qquad 2\omega - \omega^3 = 0$$

da qual

$$\omega = \pm\sqrt{2}, \qquad K = 6$$

Portanto, os lugares das raízes cruzam o eixo imaginário em $\omega = \pm\sqrt{2}$ e o valor de K nos pontos de cruzamento é 6.

5. Escolha um ponto de teste na vizinhança do eixo $j\omega$ e da origem, como indicado na Fig. 8.10, e aplique a condição do ângulo. Se um ponto de teste estiver sobre um lugar das raízes, então a soma dos três ângulos $\theta_1 + \theta_2 + \theta_3$ deve ser 180°. Se o ponto de teste não satisfizer a condição do ângulo, selecione um outro ponto de teste até que este a satisfaça. (A soma dos

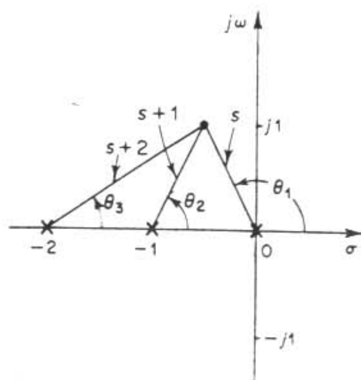


Fig. 8.10 Construção do lugar das raízes.

ângulos do ponto de teste indicará em que direção e sentido o ponto de teste deve ser movido.) Continue este processo e localize um número suficiente de pontos que satisfaça a condição do ângulo.

6. Baseado na informação obtida dos passos anteriores, desenhe um gráfico completo do lugar das raízes, como indicado na Fig. 8.11.

7. Determine um par de pólos de malha-fechada complexos conjugados dominantes tais que a relação de amortecimento ζ seja 0,5. Os pólos de malha-fechada com $\zeta = 0{,}5$ estão sobre as retas passando pela origem e fazendo ângulos de $\beta = \pm\cos^{-1}\zeta = \pm\cos^{-1}0{,}5 = \pm 60°$ com o eixo real negativo. Da Fig. 8.11, os pólos de malha-fechada possuindo $\zeta = 0{,}5$ são obtidos como segue:

$$s_1 = -0{,}33 + j0{,}58, \qquad s_2 = -0{,}33 - j0{,}58$$

O valor de K que fornece estes pólos é determinado a partir da condição de módulo como segue:

$$K = |s(s+1)(s+2)|_{s=-0,33+j0,58}$$
$$= 1{,}06$$

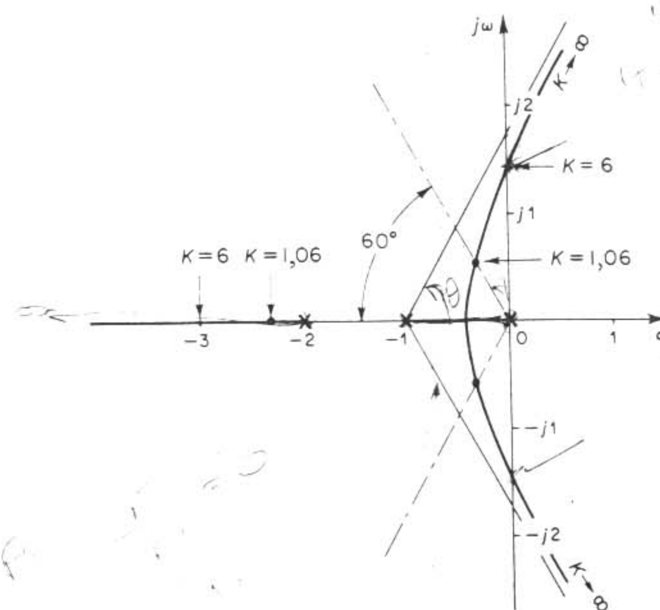


Fig. 8.11 Gráfico do lugar das raízes.

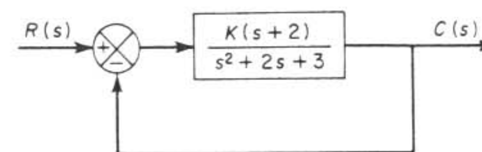


Fig. 8.12 Sistema de controle.

Usando este valor de K, é obtido o terceiro pólo em $s = -2{,}33$.

Note que, do Passo 4, pode ser verificado que para $K = 6$ os pólos em malha-fechada dominantes estão sobre o eixo imaginário em $s = \pm j\sqrt{2}$. Com este valor de K, o sistema exibirá oscilações mantidas. Para $K > 6$ os pólos de malha-fechada dominantes estão no semiplano direito do plano s, resultando em um sistema instável.

Finalmente, note que, se necessário, os lugares das raízes podem ser calibrados facilmente em termos de K pelo uso da condição do módulo. Simplesmente escolhemos um ponto sobre um lugar das raízes e medimos os módulos dos três números complexos s, $s + 1$ e $s + 2$, multiplicando estes três módulos, e o produto é igual ao valor do ganho K naquele ponto, ou

$$|s|\cdot|s + 1|\cdot|s + 2| = K$$

Exemplo 8.2 Neste exemplo, esboçaremos o gráfico do lugar das raízes de um sistema com pólos de malha-aberta complexos conjugados. Considere o sistema indicado na Fig. 8.12. Para este sistema,

$$G(s) = \frac{K(s + 2)}{s^2 + 2s + 3}, \qquad H(s) = 1$$

Verifica-se facilmente que $G(s)$ possui um par de pólos complexos conjugados em

$$s = -1 + j\sqrt{2},$$

$$s = -1 - j\sqrt{2}$$

Um procedimento típico para esboçar o gráfico do lugar das raízes é o seguinte:

1. Determine os lugares das raízes sobre o eixo real. Para qualquer ponto de teste s sobre o eixo real, a soma das contribuições angulares dos pólos complexos conjugados é 360°, como mostrado na Fig. 8.13. Portanto, a contribuição líquida dos pólos complexos conjugados é zero sobre o eixo real. A localização do lugar das raízes sobre o eixo real é determinada pelo zero de malha-aberta sobre o eixo real negativo. Um teste simples revela que a parte do eixo real negativo, entre -2 e $-\infty$, constitui uma parte do lugar das raízes. Deve ser verificado que, desde que este lugar permanece entre dois zeros (em $s = -2$ e $s = -\infty$), é na realidade uma parte de dois lugares das raízes, cada um partindo de cada um dos pólos complexos conjugados. Em outras palavras, duas raízes devem coincidir na região sobre o eixo real negativo entre -2 e $-\infty$.

Desde que há dois pólos e um zero de malha aberta, há uma assíntota que coincide com o eixo real negativo.

2. Determine o ângulo de partida dos pólos de malha-aberta complexos conjugados. A presença de um par de pólos de malha-aberta complexos conjugados requer a determinação do ângulo de partida deste pólo. É importante o conhecimento dos ângulos desde que o lugar das raízes próximo ao pólo complexo fornece informação de como o lugar das raízes originário do pólo complexo migra para o eixo real ou estende-se para a assíntota.

Referindo-se à Fig. 8.14, se escolhermos um ponto de teste e movê-lo em uma vizinhança muito próxima do pólo complexo de malha-aberta em $s = -p_1$, verificamos que a soma das contribuições angulares do pólo $s = -p_2$ e do zero $s = -z_1$ no ponto de teste pode ser considerada constante em um mesmo valor. A soma de ϕ'_1 e $-\theta'_2$ permanece a mesma conforme o ponto de teste é movido em uma vizinhança muito próxima do pólo $s = -p_1$. Se o ponto de teste estiver sobre o lugar das raízes, então a soma de ϕ'_1, $-\theta_1$ e $-\theta'_2$ deve ser $\pm 180°(2k + 1)$, onde $k = 0, 1, 2, \ldots$. Portanto, neste exemplo,

$$\phi'_1 - (\theta_1 + \theta'_2) = \pm 180°(2k + 1)$$

ou

$$\theta_1 = 180° - \theta'_2 + \phi'_1 = 180° - \theta_2 + \phi_1$$

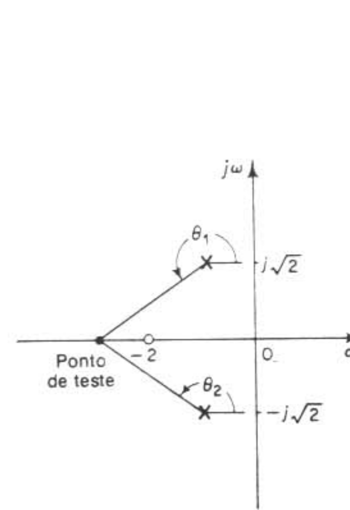


Fig. 8.13 Determinação do lugar das raízes sobre o eixo real.

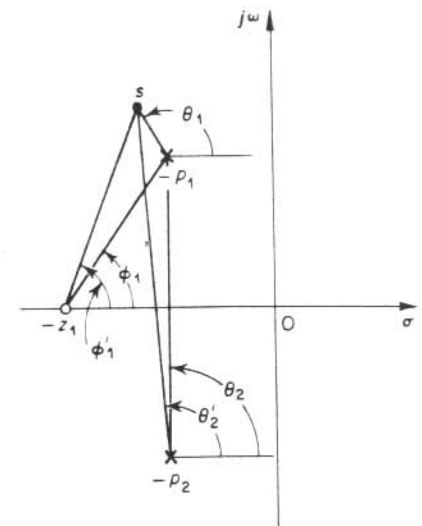


Fig. 8.14 Determinação do ângulo de partida.

O ângulo de partida é então

$$\theta_1 = 180° - \theta_2 + \phi_1 = 180° - 90° - 55° = 145°$$

Desde que o lugar das raízes é simétrico em relação ao eixo real, o ângulo de partida do pólo em $s = -p_2$ é $-145°$.

3. Determine o ponto de separação de chegada no eixo real. O ponto de separação de chegada no eixo real é aquele onde um par de ramos do lugar das raízes coincide conforme K é aumentado. O ponto de separação localiza-se entre -2 e $-\infty$ sobre o eixo real negativo e pode ser determinado como segue: Se um ponto de teste é escolhido sobre o eixo real, então a soma das contribuições angulares de um par de pólos complexos conjugados é sempre 360°. Se o ponto de teste estiver ligeiramente fora do eixo real, entretanto, então as contribuições angulares não somam 360°.

Para determinar a contribuição angular, suponha que um par de pólos complexos conjugados esteja localizado conforme a Fig. 8.15. A contribuição angular total dos pólos pode ser escrita $360° - \Delta\psi$, onde $\Delta\psi$ é obtido como segue: Da Fig. 8.15,

$$\psi_1 = \tan^{-1}\frac{a}{b}, \qquad \psi_2 = \tan^{-1}\frac{a - \delta}{b}$$

Portanto,

$$\psi_1 - \psi_2 = \tan^{-1}\frac{a}{b} - \tan^{-1}\frac{a - \delta}{b}$$

ou,

$$\tan(\psi_1 - \psi_2) = \frac{\dfrac{a}{b} - \dfrac{a - \delta}{b}}{1 + \dfrac{a(a - \delta)}{b^2}} = \frac{b\delta}{b^2 - a^2 - a\delta}$$

Se $\delta/a \ll 1$, então,

$$\psi_1 - \psi_2 \doteq \tan(\psi_1 - \psi_2) = \frac{b\delta}{a^2 + b^2}$$

A soma das contribuições angulares de um par de pólos complexos conjugados é igual a duas vezes $(\psi_1 - \psi_2)$, ou

$$2(\psi_1 - \psi_2) = \frac{2b\delta}{a^2 + b^2}$$

Portanto, em um ponto de teste próximo ao eixo real, a soma das contribuições angulares de um par de pólos complexos conjugados é dada por

$$\Delta\psi = \pm\left|\frac{2b\delta}{a^2 + b^2}\right|$$

onde o sinal mais é utilizado se os pólos estiverem à direita do ponto de teste s, enquanto o sinal negativo é utilizado se eles estiverem localizados à esquerda. Referindo-se a Fig. 8.15, verificamos que no ponto de teste s a soma dos ângulos dos três números complexos $s + p_1$, $s + p_2$, e $s + z$ é

$$(180^\circ - \Delta\phi) - \left(360^\circ - \left|\frac{2b\delta}{a^2 + b^2}\right|\right)$$

onde

$$\Delta\phi = \left|\frac{\delta}{\sigma - z}\right|$$

Se o ponto de teste está sobre o lugar das raízes, então a soma dos ângulos dos três números complexos deve ser $\pm 180^\circ$; isto é,

$$\left(180^\circ - \left|\frac{\delta}{\sigma - z}\right|\right) - \left(360^\circ - \left|\frac{2b\delta}{a^2 + b^2}\right|\right) = \pm 180^\circ$$

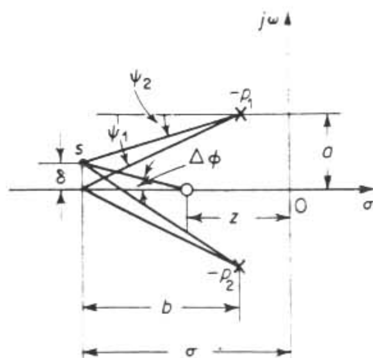


Fig. 8.15 Determinação do ponto de separação de chegada.

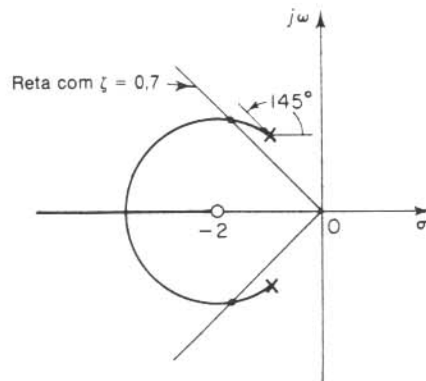


Fig. 8.16 Gráfico do lugar das raízes.

ou

$$-\left|\frac{1}{\sigma - z}\right| + \left|\frac{2b}{a^2 + b^2}\right| = 0$$

Substituindo os valores de a, b e z nesta última equação, obtemos

$$-\frac{1}{\sigma - 2} + \frac{2(\sigma - 1)}{2 + (\sigma - 1)^2} = 0$$

O valor de σ deve ser maior que dois porque o ponto de separação de chegada deve estar localizado entre -2 e $-\infty$. Resolvendo esta última equação, obtemos

$$\sigma = 3{,}73$$

Portanto, o ponto de separação de chegada dos lugares das raízes no eixo real é $(-3{,}73,\ 0)$.

4. Baseado na informação obtida dos passos anteriores, esboce um gráfico do lugar das raízes. A fim de determinar precisamente os lugares das raízes, vários pontos devem ser determinados por tentativa-e-erro entre o ponto de separação e os pólos complexos de malha-aberta. (De modo a facilitar o esboço do gráfico do lugar das raízes, devemos verificar a direção na qual o ponto de teste deve ser movido para, mentalmente, somar as variações dos ângulos devidos aos pólos e zeros.) A Fig. 8.16 mostra um gráfico completo do lugar das raízes para o sistema considerado.

O valor do ganho K em qualquer ponto do lugar das raízes pode ser determinado aplicando-se a condição do módulo. Por exemplo, o valor de K para o qual os pólos complexos conjugados de malha-fechada resultam em uma relação de amortecimento $\zeta = 0{,}7$ pode ser determinado pela localização das raízes, como indicado na Fig. 8.16, e calculado o valor K como segue:

$$K = \left|\frac{(s - 1 - j\sqrt{2})(s + 1 + j\sqrt{2})}{s + 2}\right|_{s = -1{,}67 - j1{,}70}$$
$$= 1{,}33$$

Dos exemplos anteriores, pode-se verificar claramente que é possível esboçar um gráfico do lugar das raízes razoavelmente preciso, para um dado sistema, por meio de regras simples. Em etapas preliminares de um projeto, pode-se não necessitar das posições precisas dos pólos de malha-fechada. Muitas vezes são necessárias apenas suas localizações aproximadas para se fazer uma estimativa do desempenho do sistema.

8.4 SUMÁRIO DAS REGRAS GERAIS PARA CONSTRUÇÕES DOS LUGARES DAS RAÍZES

Para um sistema complicado com muitos pólos e zeros de malha-aberta, a construção do gráfico do lugar das raízes pode parecer complicada, porém, na realidade, não é difícil se forem aplicadas as regras para construção dos lugares das raízes. Pela localização de pontos particulares e assíntotas, e pelo cálculo de ângulos de partida de pólos complexos e ângulos de chegada em zeros complexos, podemos construir a forma geral dos lugares das raízes sem dificuldade. De fato, a vantagem completa do método do lugar das raízes é poder ser realizado nos casos de sistemas de ordem superior para os quais outros métodos de determinação dos pólos em malha-fechada são extremamente trabalhosos.

Algumas das regras para construção dos lugares das raízes foram dadas na

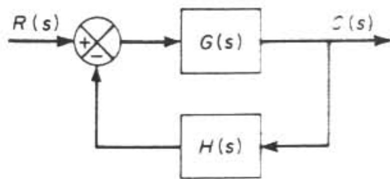


Fig. 8.17 Sistema de controle.

Seção 8.3. A finalidade desta seção é sumariar as regras gerais para construção dos lugares das raízes do sistema indicado na Fig. 8.17. Embora o método do lugar das raízes seja baseado essencialmente em uma técnica de tentativa-e-erro, o número de tentativas necessárias pode ser reduzido significativamente se usarmos estas regras.

Antes de sumariar as regras gerais, recordemos que os ângulos dos números complexos originários dos pólos de malha-aberta e dos zeros de malha-aberta até o ponto de teste s são medidos no sentido anti-horário. Por exemplo, para o sistema

$$G(s)H(s) = \frac{K(s + z_1)}{(s + p_1)(s - p_2)(s + p_3)(s + p_4)}$$

onde $-p_2$ e $-p_3$ são pólos complexos conjugados, o ângulo de $G(s)H(s)$ é

$$\underline{/G(s)H(s)} = \phi_1 - \theta_1 - \theta_2 - \theta_3 - \theta_4$$

onde ϕ_1, θ_1, θ_2, θ_3 e θ_4 são medidos no sentido anti-horário conforme indicado nas Figs. 8.18(a) e (b). O módulo de $G(s)\ H(s)$ para este sistema é

$$|G(s)H(s)| = \frac{KB_1}{A_1A_2A_3A_4}$$

onde A_1, A_2, A_3, A_4 e B_1 são os módulos dos números complexos $s + p_1$, $s + p_2$, $s + p_3$, $s + p_4$ e $s + z_1$, respectivamente, como indicado na Fig. 8.18(a).

Note que em virtude de os pólos complexos conjugados de malha-aberta e zeros complexos conjugados de malha-aberta, se houver, sempre estarem localizados simetricamente em relação ao eixo real, os lugares das raízes são sempre simétricos em relação a este eixo. Conseqüentemente, necessitamos apenas construir a metade superior dos lugares das raízes e desenhar a imagem da metade inferior, na metade inferior do plano s, como se fosse a imagem de um espelho plano posicionado sobre o eixo horizontal.

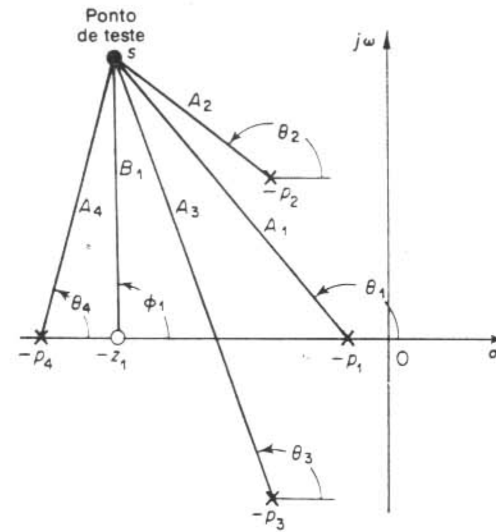
Regras gerais para construção dos lugares das raízes. Forneceremos agora um sumário das regras e procedimentos gerais para construção dos lugares das raízes do sistema mostrado na Fig. 8.17.

1. Obtenha inicialmente a equação característica

$$1 + G(s)H(s) = 0$$

e rearranje esta equação de modo que o parâmetro de interesse apareça como um fator multiplicativo na forma

$$1 + \frac{K(s + z_1)(s + z_2) \cdots (s + z_m)}{(s + p_1)(s + p_2) \cdots (s + p_n)} = 0$$



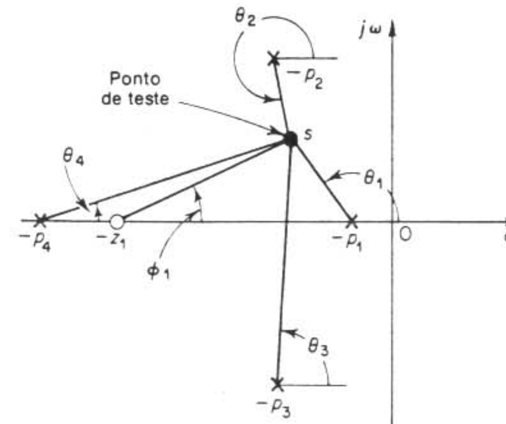


Fig. 8.18 Construção do lugar das raízes.

Nas discussões presentes, suporemos que o parâmetro de interesse é o ganho K, onde $K > 0$. (Se $K < 0$, que corresponde ao caso de realimentação positiva, a condição de ângulo deve ser modificada. Vide Problema A.8.6.) Note, entretanto, que o método ainda é aplicável a sistemas com parâmetros de interesse diferentes do ganho.

A partir da forma fatorada da função de transferência em malha-aberta, localize os pólos e zeros de malha-aberta no plano s. [Note que os zeros de malha-aberta

são os zeros de $G(s)H(s)$, enquanto os zeros de malha-fechada consistem nos zeros de $G(s)$ e nos pólos de $H(s)$.]

2. Determine os pontos de partida e os pontos de chegada dos lugares das raízes e também o número de lugares das raízes separados. Os pontos dos lugares das raízes correspondentes a $K = 0$ são os pólos de malha-aberta. Este fato pode ser verificado a partir da condição de módulo fazendo-se K tender a zero, ou

$$\lim_{K \to 0} \left| \frac{(s+z_1)(s+z_2)\cdots(s+z_m)}{(s+p_1)(s+p_2)\cdots(s+p_n)} \right| = \lim_{K \to 0} \frac{1}{K} = \infty$$

Esta última equação implica que o valor de s deve se aproximar de um dos pólos de malha-aberta. Cada lugar das raízes, portanto, origina-se em um pólo da função de transferência de malha-aberta $G(s)H(s)$. Conforme K é aumentado para infinito, cada lugar das raízes se aproxima ou de um zero da função de transferência de malha-aberta ou de um infinito no plano complexo. Isto pode ser visto como segue: Se fizermos K tender a infinito na condição do módulo, então

$$\lim_{K \to \infty} \left| \frac{(s+z_1)(s+z_2)\cdots(s+z_m)}{(s+p_1)(s+p_2)\cdots(s+p_n)} \right| = \lim_{K \to \infty} \frac{1}{K} = 0$$

Então, o valor de s deve aproximar-se de um dos zeros de malha-aberta ou de um zero de malha-aberta no infinito. (Se os zeros no infinito forem incluídos na contagem, $G(s)H(s)$ possui o mesmo número de pólos e zeros.)

Um gráfico do lugar das raízes possuirá tantos ramos quantas forem as raízes da equação característica. Desde que o número de pólos em malha-aberta geralmente excede o número de zeros, o número de ramos é igual ao dos pólos. Se o número de pólos em malha-fechada é o mesmo do número de pólos de malha-aberta, então o número de ramos individuais do lugar das raízes terminando em um zero finito de malha-aberta é igual ao número m de zeros de malha-aberta. Os restantes $n - m$ ramos terminam no infinito ($n - m$ zeros implícitos no infinito) ao longo das assíntotas.

É importante notar, entretanto, que se considerarmos um problema puramente matemático, então, pode-se fazer com que o número de pólos de malha-fechada se iguale ao número de zeros de malha-aberta ao invés do número de pólos de malha-aberta. Neste caso, o número de ramos do lugar das raízes é igual ao número de zeros de malha-aberta. Por exemplo, considere a seguinte equação polinomial:

$$s^2 + s + 1 = 0$$

Esta equação pode ser reescrita

$$1 + \frac{s^2}{s+1} = 0$$

Então a função de transferência $s^2/(s + 1)$ pode ser considerada a função de transferência de malha-aberta. Ela possui dois zeros e um pólo. Portanto, o número de zeros finitos é maior do que o número de pólos finitos. O número de ramos do lugar das raízes é igual àquele dos zeros de malha-aberta.

Se incluirmos os pólos e zeros no infinito, o número de pólos de malha-aberta é igual ao número de zeros de malha-aberta. Conseqüentemente, podemos sempre estabelecer que os lugares das raízes iniciam-se nos pólos de $G(s)H(s)$ e terminam nos zeros de $G(s)H(s)$, conforme K aumenta de zero a infinito, onde os pólos e zeros incluem tanto aqueles no plano s finito como aqueles no infinito.

3. Determine os lugares das raízes sobre o eixo real. Os lugares das raízes no eixo real são determinados pelos pólos e zeros de malha-aberta nele contidos. Os pólos e zeros complexos conjugados da função de transferência de malha-aberta não afetam a localização dos lugares das raízes sobre o eixo real porque a contribuição angular de um par de pólos ou zeros complexos conjugados é 360° no eixo real. Cada parte do lugar das raízes sobre o eixo real estende-se em um intervalo desde um pólo ou zero até um outro pólo ou zero. Na construção dos lugares das raízes sobre o eixo real, escolhemos um ponto de teste sobre o mesmo eixo. Se o número total de pólos reais e zeros reais à direita deste ponto de teste for ímpar, então este ponto pertence a um lugar das raízes. O lugar das raízes e sua forma complementar, alternam segmentos ao longo do eixo real.

4. Determine as assíntotas dos lugares das raízes. Se o ponto de teste s for localizado distante da origem, então o ângulo de cada número complexo pode ser considerado o mesmo. Um zero de malha-aberta e um pólo de malha-aberta cancelam os efeitos um do outro. Portanto, os lugares das raízes para valores muito grandes de s devem ser assintóticos com retas cujos ângulos (inclinações) são dados por

$$\text{Ângulo da assíntota} = \frac{\pm 180°(2k+1)}{n-m} \qquad (k = 0, 1, 2, \ldots)$$

onde

n = número de pólos finitos de $G(s)H(s)$
m = número de zeros finitos de $G(s)H(s)$

Neste caso, $k = 0$ corresponde a assíntotas com os menores ângulos em relação ao eixo real. Embora k possua um número infinito de valores, conforme k aumenta o ângulo se repete e o número de assíntotas distintas é $n - m$.

Todas as assíntotas se interceptam entre si sobre o eixo real. O ponto no qual elas se interceptam é obtido como segue: Se tanto o numerador como o denominador da função de transferência de malha-aberta forem expandidos, o resultado é

$$G(s)H(s) = \frac{K[s^m + (z_1 + z_2 + \cdots + z_m)s^{m-1} + \cdots + z_1 z_2 \cdots z_m]}{s^n + (p_1 + p_2 + \cdots + p_n)s^{n-1} + \cdots + p_1 p_2 \cdots p_n}$$

Se um ponto de teste está localizado muito distante da origem, então $G(s)H(s)$ pode ser escrita

$$G(s)H(s) = \\ = \frac{K}{s^{n-m} + [(p_1 + p_2 + \cdots + p_n) - (z_1 + z_2 + \cdots + z_m)]s^{n-m-1} + \cdots}$$

Desde que a equação característica é

$$G(s)H(s) = -1$$

ela pode ser escrita

$$s^{n-m} + [(p_1 + p_2 + \cdots + p_n) - \\ - (z_1 + z_2 + \cdots + z_m)]s^{n-m-1} + \cdots = -K \qquad (8.12)$$

Para um valor grande de s, a Eq. (8.12) pode ser aproximada por

$$\left[s + \frac{(p_1 + p_2 + \cdots + p_n) - (z_1 + z_2 + \cdots + z_m)}{n - m}\right]^{n-m} = 0$$

Se a abscissa da intersecção das assíntotas e o eixo real for indicada por $-\sigma_a$, então

$$-\sigma_a = -\frac{(p_1 + p_2 + \cdots + p_n) - (z_1 + z_2 + \cdots + z_m)}{n - m} \qquad (8.13)$$

Como todos os pólos e zeros complexos ocorrem em pares conjugados, $-\sigma_a$ é sempre um número real. Note que $-\sigma_a$ é o centróide dos pólos e zeros de $G(s)H(s)$, onde o centróide é calculado considerando-se que todos os pólos possuem massa $+1$ e todos os zeros possuem massa -1. Uma vez determinada a intersecção das assíntotas e o eixo real, elas podem ser realmente desenhadas no plano complexo.

5. Determine os pontos de separação de partida e chegada em relação ao eixo real. Devido à simetria conjugada dos lugares das raízes, os pontos de separação de partida e de chegada ou permanecem no eixo real, ou ocorrem em pares complexos conjugados.

Se um lugar das raízes estiver entre dois pólos de malha-aberta adjacentes sobre o eixo real, então existe pelo menos um ponto de separação de partida entre os dois pólos. Analogamente, se existir um lugar das raízes entre dois zeros adjacentes (um zero pode estar localizado em $-\infty$) sobre o eixo real, então sempre existirá pelo menos um ponto de separação de chegada entre os dois zeros. Se existir um lugar das raízes entre um pólo e um zero (finito ou infinito) de malha-aberta sobre o eixo real, então não podem existir pontos de separação de partida ou de chegada, ou então, lá existirá tanto pontos de separação de partida como de chegada.

Se a equação característica é dada por

$$A(s) - KB(s) = 0$$

os pontos de separação de partida e de chegada podem ser determinados das raízes de

$$\frac{dK}{ds} = -\frac{A'(s)B(s) - A(s)B'(s)}{B^2(s)} = 0$$

onde o apóstrofo indica diferenciação em relação a s. Se o valor de K correspondente a uma raiz $s = s_1$ de $dK/ds = 0$ é positivo, o ponto $s = s_1$ é um ponto de separação de partida ou chegada real. (Desde que se supõe que o valor de K é não-negativo, se o valor de K é obtido negativo, então o ponto $s = s_1$ não é um ponto de separação de chegada ou partida.) Note que esta abordagem pode, obviamente, ser utilizada quando houver pólos e/ou zeros complexos.

6. Determine os ângulos de partida (ou ângulos de chegada) dos lugares das raízes dos pólos complexos (ou de zeros complexos). De modo a esboçar os lugares das raízes com razoável precisão, devemos determinar as direções dos lugares das raízes próximas aos pólos e zeros complexos. Se for escolhido um ponto de teste e este for movido em uma vizinhança muito próxima de um pólo complexo (ou zero complexo), pode-se considerar que a soma das contribuições angulares de todos os outros pólos e zeros permanece a mesma. Portanto, o ângulo de partida (ou o ângulo de chegada) do lugar das raízes de um pólo complexo (ou de um zero complexo) pode ser determinado subtraindo-se de 180° a soma de todos os ângulos dos

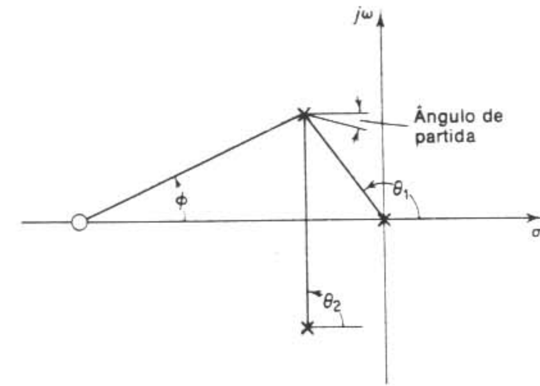


Fig. 8.19 Construção do lugar das raízes. [Ângulo de partida = 180° − ($\theta_1 + \theta_2 - \phi_1$)]

números complexos desde os outros pólos e zeros até o pólo complexo (ou zero complexo) em questão, incluindo-se os sinais apropriados. O ângulo de partida é indicado na Fig. 8.19.

7. Determine os pontos onde os lugares das raízes cruzam o eixo imaginário. Os pontos onde os lugares das raízes interceptam o eixo $j\omega$ podem ser facilmente determinados pelo uso do critério de estabilidade de Routh através de uma abordagem de tentativa-e-erro ou pela substituição de $s = j\omega$ na equação característica, igualando a zero tanto a parte real quanto a parte imaginária e resolvendo para obter-se ω e K. O valor de ω assim determinado fornece a freqüência na qual o lugar das raízes cruza o eixo imaginário, e o valor de K corresponde a um ganho crítico para estabilidade.

8. A equação característica do sistema cuja função de transferência de malha-aberta é

$$G(s)H(s) = \frac{K(s^m + b_1 s^{m-1} + \cdots + b_m)}{s^n + a_1 s^{n-1} + \cdots + a_n} \qquad (n \geq m)$$

é uma equação algébrica em s de grau n. Se a ordem do numerador de $G(s)H(s)$ é inferior àquela do denominador por dois ou mais (o que significa que há dois ou mais zeros no infinito), então o coeficiente a_1 é a soma negativa das raízes da equação e é independente de K. Neste caso, se alguma das raízes move-se para a esquerda sobre o lugar das raízes conforme K é aumentado, então as outras raízes devem mover-se para a direita, conforme K é aumentado. Esta informação é útil na determinação da forma geral dos lugares das raízes.

9. Determine os lugares das raízes na vizinhança do eixo $j\omega$ e da origem. A parte mais importante dos lugares das raízes não está nem sobre o eixo real nem nas assíntotas mas na parte da vizinhança do eixo $j\omega$ e da origem. A forma dos lugares das raízes nesta importante região do plano s deve ser obtida com precisão suficiente.

Na construção de um gráfico do lugar das raízes, necessitamos apenas de um transferidor e uma régua. Um dispositivo denominado Espírula, consistindo em um transferidor e uma régua girante a ele acoplada, é particularmente conveniente para verificação da condição de ângulo em pontos de teste. (Uma escala logarítmica espiral impressa na Espírula pode ser utilizada na determinação do ganho K a partir da condição do módulo.)

Alguns poucos comentários sobre os gráficos do lugar das raízes. Qualquer ponto sobre os lugares das raízes é um pólo de malha-fechada possível. Um ponto particular será um pólo de malha-fechada quando o valor do K satisfizer a condição do módulo. Portanto, a condição do módulo nos possibilita determinar o valor do ganho K para qualquer localização específica de uma raiz sobre o lugar das raízes. (Se necessário, os lugares das raízes podem ser graduados em termos de K. Os lugares das raízes são contínuos com K.)

Se o ganho K da função de transferência de malha-aberta é dado no problema, então, aplicando-se a condição de módulo podem-se determinar as posições corretas dos pólos de malha-fechada, para o K especificado, sobre cada ramo do lugar das raízes através de uma abordagem de tentativa-e-erro.

Uma vez determinados os pólos de malha-fechada dominantes (ou os pólos de malha-fechada mais próximos do eixo $j\omega$) pelo método do lugar das raízes, então os pólos de malha-fechada restantes podem ser obtidos dividindo-se a equação característica pelos fatores correspondentes aos pólos de malha-fechada dominantes. O resto corresponde aos pólos de malha-fechada restantes. Muitas vezes esta divisão pode não ser efetuada exatamente. Isto é inevitável devido às imprecisões introduzidas pela análise gráfica.

Como um comentário final, note que uma pequena variação na configuração de pólos e zeros pode acarretar variações significativas na configuração do lugar das raízes. A Fig. 8.20 demonstra o fato de que uma pequena modificação na posição do zero resultará em uma configuração do lugar das raízes bastante diferente.

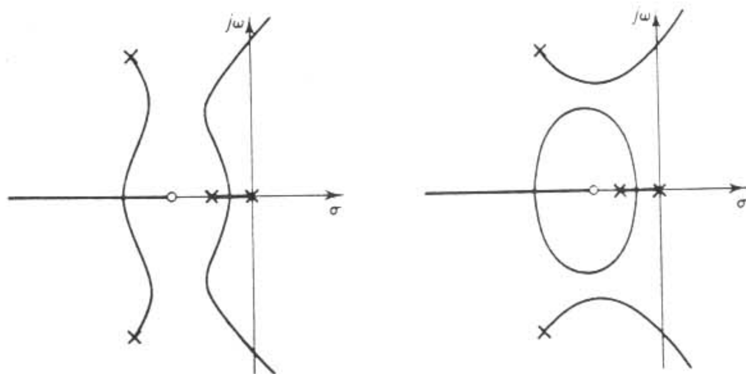
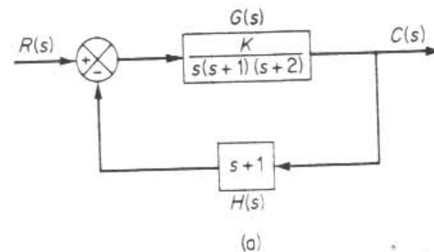


Fig. 8.20 Gráficos do lugar das raízes.

Cancelamento de pólos de $G(s)$ com zeros de $H(s)$. É importante notar que se o denominador de $G(s)$ e o numerador de $H(s)$ envolvem fatores comuns, então os pólos e zeros de malha-aberta correspondentes se cancelarão entre si, reduzindo o grau da equação característica de um ou mais. Por exemplo, considere o sistema mostrado na Fig. 8.21(a). (Este sistema possui realimentação de velocidade.) A função de transferência de malha-fechada $C(s)/R(s)$ é

$$\frac{C(s)}{R(s)} = \frac{K}{s(s+1)(s+2) + K(s+1)}$$



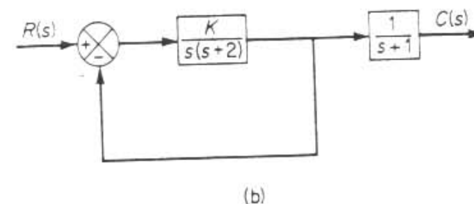


Fig. 8.21 Sistemas de controle.

A equação característica é

$$[s(s+2) + K](s+1) = 0$$

Devido ao cancelamento dos termos $(s+1)$ que aparecem em $G(s)$ e $H(s)$, entretanto, temos

$$1 + G(s)H(s) = 1 + \frac{K(s+1)}{s(s+1)(s+2)} = \frac{s(s+2) + K}{s(s+2)}$$

A equação característica reduzida é

$$s(s+2) + K = 0$$

O gráfico do lugar das raízes de $G(s)H(s)$ não mostra todas as raízes da equação característica, apenas as raízes da equação reduzida.

Para obter o conjunto completo de pólos de malha-fechada, devemos adicionar os pólos cancelados de $G(s)H(s)$ aos pólos de malha-fechada obtidos do gráfico do lugar das raízes de $G(s)H(s)$. O fato importante a lembrar é que o pólo cancelado de $G(s)H(s)$ é um pólo de malha-fechada do sistema, conforme visto na Fig. 8.21(b).

8.5 ANÁLISE DE SISTEMAS DE CONTROLE PELO MÉTODO DO LUGAR DAS RAÍZES

Nesta seção, examinaremos inicialmente os efeitos de um zero na configuração dos lugares das raízes de um sistema de segunda-ordem. Posteriormente, compararemos os efeitos do controle derivativo e realimentação de velocidade no desempenho de um servomecanismo de posição. A seguir apresentaremos uma análise de um sistema condicionalmente estável e sistemas de fase não mínima.

Finalmente, voltaremos nossa atenção à análise das variações de dois parâmetros do sistema.

Efeitos de um zero nos lugares das raízes de sistemas de segunda-ordem. Considere o sistema indicado na Fig. 8.22(a). A função de transferência de malha-aberta $G(s)H(s)$ é

$$G(s)H(s) = \frac{K(s - z)}{(s + p_1)(s - p_2)} \qquad (K \geq 0,\ p_1 > p_2 > 0,\ z > 0)$$

Examinaremos os lugares das raízes deste sistema para os três casos seguintes:

Caso 1: $z > p_1 > p_2$

Caso 2: $p_1 > p_2 > z$

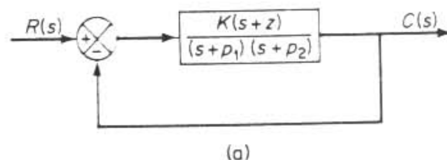
Caso 3: $p_1 > z > p_2$

Para o Caso 1, os pontos de separação podem ser obtidos como segue: Desde que a equação característica para este sistema é

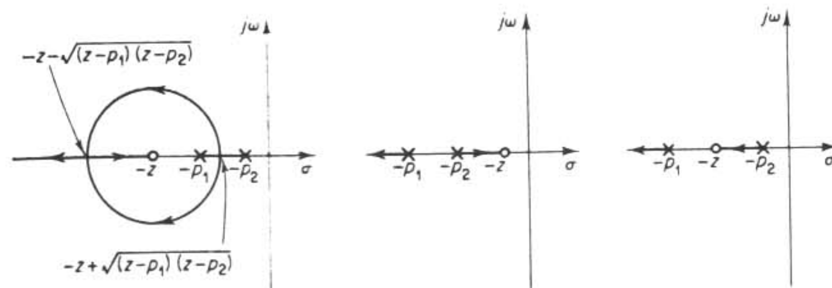
$$(s + p_1)(s + p_2) + K(s - z) = 0$$

ou

$$K = -\frac{(s + p_1)(s - p_2)}{s + z}$$



(a)



(b)

Fig. 8.22 (a) Sistema de controle; (b) gráficos do lugar das raízes.

as raízes de $dK/ds = 0$ podem ser obtidas de

$$(s + p_1)(s + p_2) - (2s + p_1 + p_2)(s + z) = 0$$

como segue:

$$s = -z \pm \sqrt{(z - p_1)(z - p_2)} \tag{8.14}$$

Desde que $z > p_1 > p_2$, a quantidade dentro da raiz quadrada é positiva. O valor de K correspondente a $s = -z + \sqrt{(z - p_1)(z - p_2)}$ é

$$K = (\sqrt{z - p_1} - \sqrt{z - p_2})^2 > 0$$

Analogamente, o valor de K correspondente a $s = -z - \sqrt{(z - p_1)(z - p_2)}$ é

$$K = (\sqrt{z - p_1} + \sqrt{z - p_2})^2 > 0$$

Conseqüentemente, tanto $s = -z + \sqrt{(z - p_1)(z - p_2)}$ como $s = -z - \sqrt{(z - p_1)(z - p_2)}$ são pontos de separação.

Para o Caso 2, embora o argumento da raiz quadrada da Eq. (8.14) seja positivo, e portanto os valores de s dados pela Eq. (8.14) sejam reais, os valores correspondentes de K tornam-se negativos desde que

$$K = -(\sqrt{p_1 - z} \pm \sqrt{p_2 - z})^2 < 0$$

Isto significa que para o Caso 2 os pontos dados pela Eq. (8.14) não são pontos de separação, e, portanto, não existem pontos de separação.

Para o Caso 3, o argumento da raiz quadrada da Eq. (8.14) torna-se negativo, significando que os pontos dados pela Eq. (8.14) são complexos conjugados. Desde que os pontos de separação de chegada e de partida, se existirem, devem estar sobre o eixo real no presente exemplo, os pontos dados pela Eq. (8.14) não correspondem a pontos de separação. Portanto, não existem pontos de separação no Caso 3, e os lugares das raízes são simplesmente dois segmentos do eixo real negativo.

A Fig. 8.22(b) mostra os gráficos do lugar das raízes correspondentes aos três casos considerados.

Comparação dos efeitos do controle derivativo e realimentação de velocidade (realimentação tacométrica) no desempenho de servomecanismos de posição. O Sistema I mostrado na Fig. 8.23 é um servomecanismo de posição. (A saída é posição.) O Sistema II mostrado na Fig. 8.23 é um servomecanismo de posição utilizando ação de controle proporcional-mais-derivativo. O Sistema III mostrado na Fig. 8.23 é um servomecanismo de posição utilizando realimentação em velocidade ou realimentação tacométrica. Vamos comparar os méritos relativos do controle derivativo e da realimentação de velocidade.

O gráfico do lugar das raízes para o Sistema I é mostrado na Fig. 8.24(a). Os pólos de malha-fechada estão localizados em $s = -0{,}1 \pm j0{,}995$.

A função de transferência de malha-aberta do Sistema II é

$$G_{II}(s)H_{II}(s) = \frac{5(1 + 0{,}8s)}{s(5s + 1)}$$

A função de transferência de malha-aberta do Sistema III é

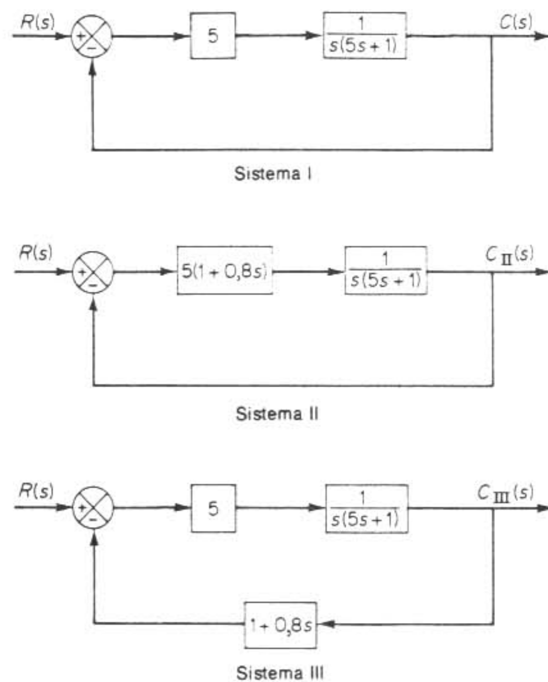


Fig. 8.23 Servomecanismos de posição.

$$G_{III}(s)H_{III}(s) = \frac{5(1 + 0{,}8s)}{s(5s + 1)}$$

Portanto os Sistemas II e III possuem funções de transferência de malha-aberta idênticas. (Ambos os sistemas têm os mesmos pólos e zeros de malha-aberta.) Os gráficos do lugar das raízes dos Sistemas II e III são portanto idênticos e são fornecidos nas Figs. 8.24 (b) e (c), respectivamente.

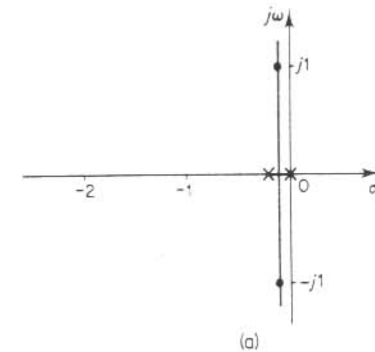
Note, entretanto, que a função de transferência de malha-fechada do Sistema II é evidentemente diferente daquela do Sistema III. O Sistema II possui dois pólos de malha-fechada e um zero de malha-fechada finito, enquanto o Sistema III possui dois pólos de malha-fechada e não possui zero de malha-fechada finito. (A realimentação de velocidade, ou realimentação tacométrica, possui um zero de malha-aberta mas nenhum zero de malha-fechada.) Os pólos de malha-fechada dos Sistemas II e III são $s = -0{,}5 \pm j0{,}866$.

A função de transferência de malha-fechada do Sistema II é

$$\frac{C_{II}(s)}{R(s)} = \frac{1 + 0{,}8s}{(s + 0{,}5 + j0{,}866)(s + 0{,}5 - j0{,}866)}$$

Para uma entrada impulsiva-unitária,

$$C_{II}(s) = \frac{0{,}4 + j0{,}346}{s + 0{,}5 + j0{,}866} + \frac{0{,}4 - j0{,}346}{s + 0{,}5 - j0{,}866}$$



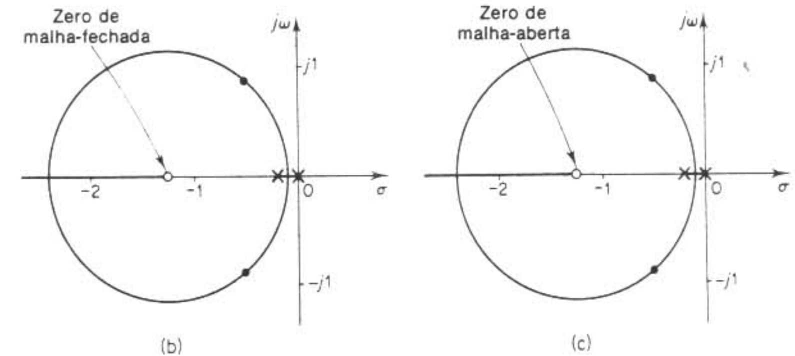


Fig. 8.24 Gráficos do lugar das raízes dos sistemas indicados na Fig. 8.23.

O resíduo do pólo de malha-fechada $s = -0{,}5 - j0{,}866$ é $0{,}4 + j0{,}346$, e aquele do pólo de malha-fechada $s = -0{,}5 + j0{,}866$ é $0{,}4 - j0{,}346$. A transformada inversa de Laplace de $C_{II}(s)$ resulta

$$c_{II}(t)_{\text{impulso}} = e^{-0{,}5t}(0{,}8 \cos 0{,}866t + 0{,}693 \operatorname{sen} 0{,}866t) \qquad (t \geq 0)$$

A função de transferência de malha-fechada do Sistema III é

$$\frac{C_{III}(s)}{R(s)} = \frac{1}{(s + 0{,}5 + j0{,}866)(s + 0{,}5 - j0{,}866)}$$

Para uma entrada impulsiva-unitária,

$$C_{III}(s) = \frac{j0{,}577}{s + 0{,}5 + j0{,}866} + \frac{-j0{,}577}{s + 0{,}5 - j0{,}866}$$

O resíduo no pólo de malha-fechada $s = -0{,}5 - j0{,}866$ é $j0{,}577$, e aquele no pólo de malha-fechada $s = -0{,}5 + j0{,}866$ é $-j0{,}577$. A transformada de Laplace inversa de $C_{III}(s)$ resulta

$$c_{III}(t)_{\text{impulso}} = 1{,}155e^{-0{,}5t} \operatorname{sen} 0{,}866t \qquad (t \geq 0)$$

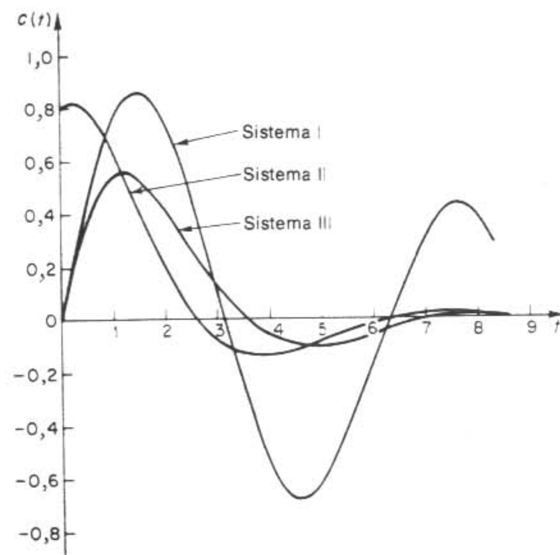


Fig. 8.25 Curvas de resposta ao impulso unitário para os Sistemas I, II e III indicados na Fig. 8.23.

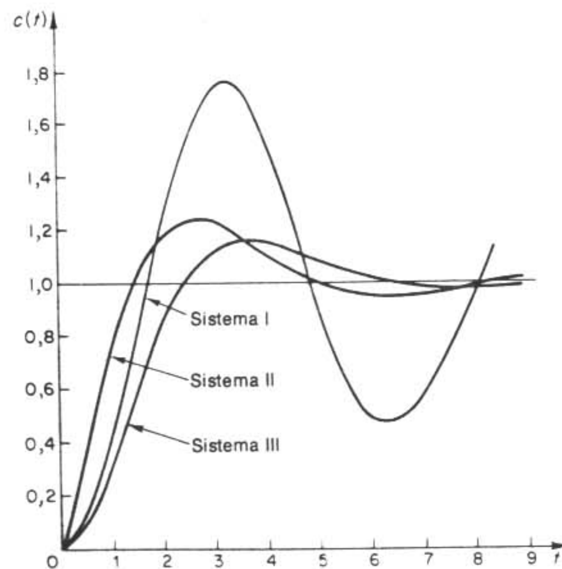


Fig. 8.26 Curvas de resposta ao degrau unitário para os Sistemas I, II e III indicados na Fig. 8.23.

As respostas ao impulso-unitário dos Sistemas II e III são evidentemente diferentes porque os resíduos no mesmo pólo são diferentes para os dois sistemas. (Lembre-se que os resíduos dependem tanto dos pólos como dos zeros de malha-fechada.) A Fig. 8.25 mostra as curvas de resposta ao impulso-unitário para os três sistemas.

Note que a resposta ao degrau unitário pode ser obtida, ou diretamente, ou pela integração da resposta ao impulso unitário. Por exemplo, para o Sistema III, a resposta do degrau unitário é obtida como segue:

$$c_{\text{III}}(t)_{\text{degrau}} = \int_0^t c_{\text{III}}(t)_{\text{impulso}}\, dt$$
$$= \int_0^t 1{,}155 e^{-0{,}5t} \operatorname{sen} 0{,}866t\, dt$$
$$= 1 - e^{-0{,}5t}(\cos 0{,}866t + 0{,}577 \operatorname{sen} 0{,}866t)$$

A Fig. 8.26 mostra as curvas de resposta ao degrau unitário para os três sistemas. O sistema utilizando ação de controle proporcional-mais-derivativa apresenta o menor tempo de subida. O sistema com realimentação de velocidade possui a mínima sobrelevação, ou a melhor estabilidade relativa dos três sistemas.

A Fig. 8.27 mostra as curvas de resposta a rampa-unitária para os Sistemas II e III. O Sistema II possui a vantagem da resposta mais rápida e menor erro em regime permanente para a entrada em rampa.

A principal razão de o sistema que utiliza ação de controle proporcional-mais-derivativa possuir características de resposta superiores é que o controle derivativo responde à taxa de variação do sinal de erro e pode produzir ação corretiva antes de a amplitude do erro tornar-se grande.

Note que a saída do Sistema III é a saída do Sistema II atrasada por um termo atrasador de primeira ordem $1/(1 + 0{,}8s)$. A Fig. 8.28 mostra a relação entre a saída do Sistema II e a saída do Sistema III.

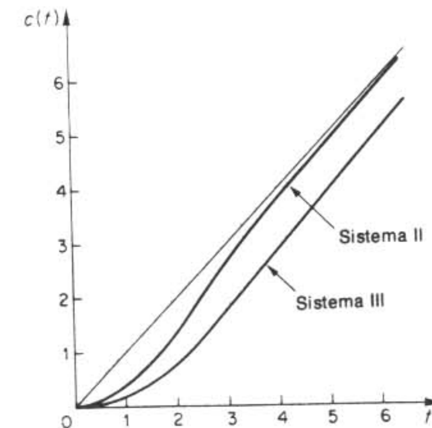


Fig. 8.27 Curvas de resposta à rampa unitária para os Sistemas II e III indicados na Fig. 8.23.

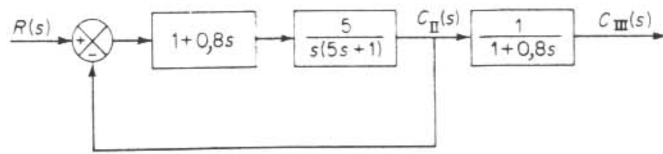


Fig. 8.28 Diagrama de blocos indicando a relação entre as saídas dos Sistemas II e III, indicados na Fig. 8.23.

Sistemas condicionalmente estáveis. Considere o sistema indicado na Fig. 8.29(a). Os lugares das raízes para este sistema podem ser construídos aplicando-se as regras e procedimentos gerais para construção dos lugares das raízes. Um gráfico do lugar das raízes para este sistema é fornecido na Fig. 8.29(b). Pode ser visto que este sistema é estável apenas em intervalos limitados do valor de K; isto é, $14 > K > 0$ e $195 > K > 64$. O sistema torna-se instável para $64 > K > 14$ e $K > 195$. Se K possuir um valor correspondente à operação instável, o sistema pode deixar de funcionar ou tornar-se não linear devido à existência de uma não linearidade por saturação. Um sistema deste tipo é denominado *condicionalmente estável*.

Na prática, sistemas condicionalmente estáveis não são desejáveis. A estabilidade condicional é perigosa, porém ocorre em alguns sistemas, em particular um sistema que possui um caminho instável no ramo direto. Este caminho pode ocorrer se o sistema possuir um pequeno laço. É aconselhável evitar esta estabilidade condicional desde que, se o ganho cair além do valor crítico por alguma razão, o sistema torna-se instável. Note que a adição de uma rede de compensação apropriada eliminará a estabilidade condicional. [Uma adição de um zero causará o deslocamento dos lugares das raízes para a esquerda. (Vide Seção 10.3.) Portanto a estabilidade condicional pode ser eliminada pela adição de uma compensação apropriada.]

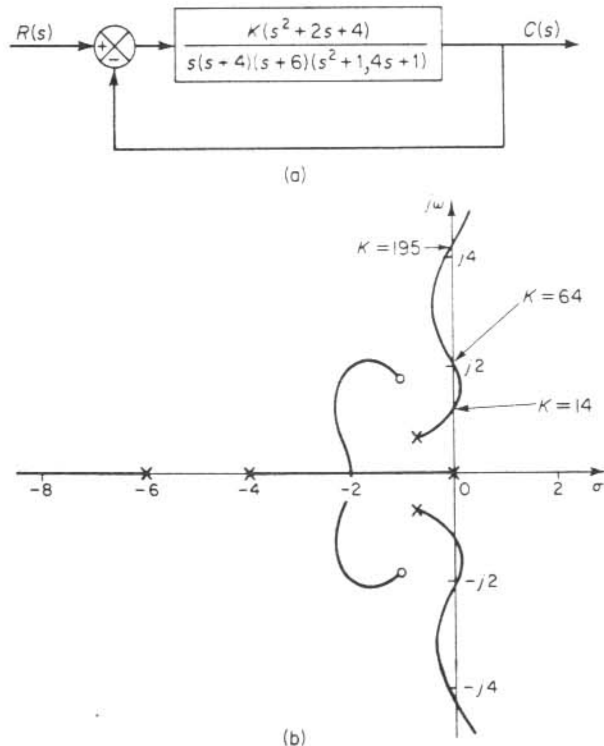


Fig. 8.29 (a) Sistema de controle; (b) gráfico do lugar das raízes.

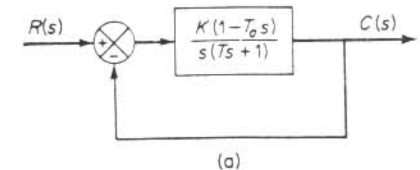
Sistemas de fase não mínima. Se todos os pólos e zeros de um sistema estiverem no semiplano esquerdo do plano s, então o sistema é chamado de fase mínima. Se um sistema possuir pelo menos um pólo ou um zero no semiplano direito do plano s, então o sistema é denominado sistema de fase não-mínima. O termo "fase não-mínima" provém da característica de deslocamento de fase de tais sistemas quando sujeitos a entradas senoidais. (Vide Seção 9.2.)

Considere o sistema indicado na Fig. 8.30(a). Neste sistema

$$G(s) = \frac{K(1 - T_a s)}{s(Ts + 1)} \qquad (T_a > 0), \qquad H(s) = 1$$

Este é um sistema de fase não-mínima desde que há um zero no semiplano direito do plano s. Para este sistema, a condição de ângulo resulta

$$\angle G(s) = \angle -\frac{K(T_a s - 1)}{s(Ts + 1)} = \angle \frac{K(T_a s - 1)}{s(Ts + 1)} + 180^\circ$$
$$= \pm 180^\circ(2k + 1) \qquad (k = 0, 1, 2, \ldots)$$



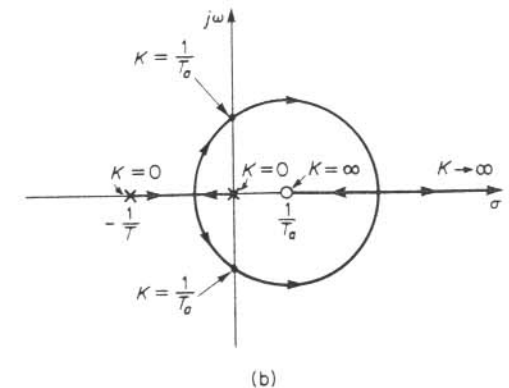


Fig. 8.30 (a) Sistema realimentado; (b) gráfico do lugar das raízes.

ou

$$\angle \frac{K(T_a s - 1)}{s(Ts + 1)} = 0° \qquad (8.15)$$

Os lugares das raízes podem ser obtidos da Eq. (8.15). A Fig. 8.30(b) mostra um gráfico do lugar das raízes para este sistema. Deste gráfico, verificamos que o sistema é estável se o ganho K é menor do que $1/T_a$.

Sistemas com atraso de transporte ou tempo morto. A Fig. 8.31 mostra um sistema térmico no qual é circulado ar quente a fim de manter constante a temperatura de uma câmara. Neste sistema o elemento de medida é colocado no retorno do fluxo a uma distância L do forno, sendo a velocidade do ar v e o tempo decorrido, antes de qualquer variação na temperatura do forno ser sentida pelo termômetro, considerado T. Este atraso na medida, atraso na ação do controlador, ou atraso na operação do atuador etc. é chamado *tempo morto* ou *atraso de transporte*. O tempo morto existe na maioria dos sistemas de controle de processos.

A entrada $x(t)$ e a saída $y(t)$ de um elemento com tempo morto são relacionadas por

$$y(t) = x(t - T)$$

onde T é o tempo morto. A função de transferência do tempo morto é dada por

$$\text{Função de transferência do tempo morto ou atraso de transporte} = \frac{\mathscr{L}[x(t-T)1(t-T)]}{\mathscr{L}[x(t)1(t)]} = \frac{X(s)e^{-Ts}}{X(s)} = e^{-Ts}$$

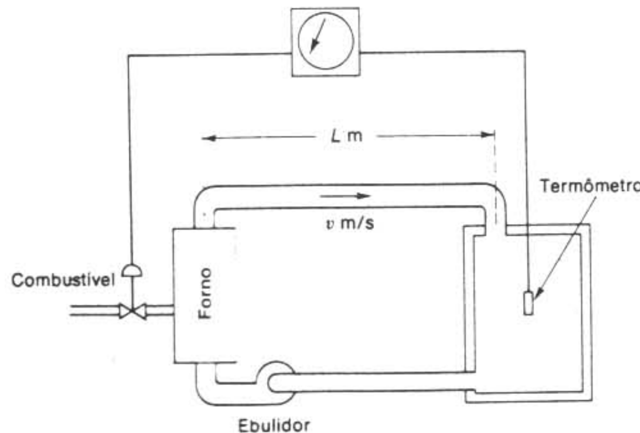


Fig. 8.31 Sistema térmico.

Suponha que a função de transferência do ramo direto deste sistema térmico possa ser aproximada por

$$G(s) = \frac{Ke^{-Ts}}{s + 1}$$

conforme indicada na Fig. 8.32. Vamos construir um gráfico do lugar das raízes para este sistema. A equação característica para este sistema de malha-fechada é

$$1 + \frac{Ke^{-Ts}}{s + 1} = 0 \qquad (8.16)$$

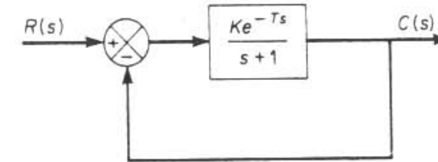


Fig. 8.32 Diagrama de blocos do sistema indicado na Fig. 8.31.

Da Eq. (8.16), obtemos

$$\frac{Ke^{-Ts}}{s + 1} = -1$$

Portanto, a condição de ângulo resulta

$$\angle \frac{Ke^{-Ts}}{s+1} = \angle e^{-Ts} - \angle s + 1 = \pm 180°(2k + 1) \qquad (k = 0, 1, 2, \ldots) \quad (8.17)$$

Para determinar o ângulo de e^{-Ts}, vamos escrever $s = \sigma + j\omega$. Obtemos então

$$e^{-Ts} = e^{-T\sigma - j\omega T}$$

Desde que $e^{-T\sigma}$ é um número real, o ângulo de $e^{-T\sigma}$ é zero. Portanto,

$$\begin{aligned} \angle e^{-Ts} &= \angle e^{-j\omega T} = \angle \cos \omega T - j \operatorname{sen} \omega T \\ &= -\omega T \qquad \text{(radianos)} \\ &= -57{,}3\omega T \qquad \text{(graus)} \end{aligned}$$

A condição do ângulo, Eq. (8.17), torna-se

$$-57{,}3\omega T - \angle s + 1 = \pm 180°(2k + 1)$$

Desde que T é uma constante dada, o ângulo de e^{-Ts} é uma função apenas de ω.

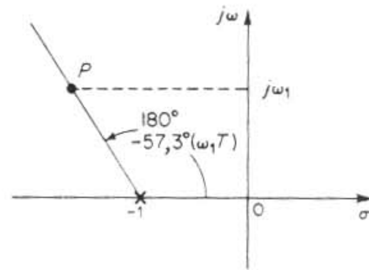
Determinemos agora a contribuição angular devido ao termo e^{-Ts}. Para $k = 0$, a condição de ângulo pode ser escrita

$$\angle s + 1 = \mp 180° - 57{,}3°\omega T \qquad (8.18)$$

Desde que a contribuição angular de e^{-Ts} é zero para $\omega = 0$, o eixo real desde -1 até $-\infty$ constitui uma parte do lugar das raízes. Suponha agora um valor ω_1 para ω e calcule $57{,}3^\circ\omega_1 T$. No ponto -1 sobre o eixo real negativo, desenhe uma reta que faz um ângulo de $180^\circ - 57{,}3^\circ\omega_1 T$ com o eixo real. Determine a intersecção desta reta com a reta horizontal $\omega = \omega_1$. Esta intersecção, ponto P na Fig. 8.33(a), é um ponto que satisfaz a Eq. (8.18) e portanto está sobre o lugar das raízes. Continuando o mesmo processo, obtemos o gráfico do lugar das raízes indicado na Fig. 8.33(b).

Note que conforme s tende a menos infinito, a função de transferência de malha-aberta

$$\frac{Ke^{-Ts}}{s+1}$$



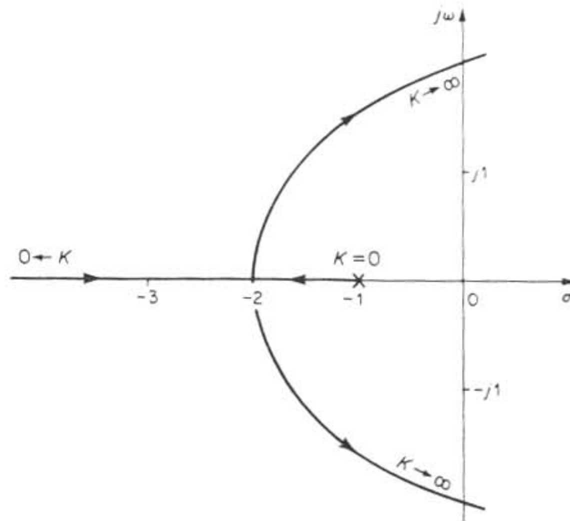


Fig. 8.33 (a) Construção do lugar das raízes; (b) gráfico do lugar das raízes.

tende a menos infinito, desde que

$$\lim_{s=-\infty} \frac{Ke^{-Ts}}{s-1} = \left.\frac{\dfrac{d}{ds}(Ke^{-Ts})}{\dfrac{d}{ds}(s+1)}\right|_{s=-\infty}$$
$$= -KTe^{-Ts}\big|_{s=-\infty}$$
$$= -\infty$$

Portanto, $s = -\infty$ é um pólo da função de transferência de malha-aberta. Conseqüentemente, os lugares das raízes têm início em $s = -1$ ou $s = -\infty$ e terminam em $s = \infty$, conforme K aumenta de zero a infinito. Desde que o segundo membro da condição de ângulo dada pela Eq. (8.17) possui um número infinito de valores, há um número infinito de lugares das raízes, conforme o valor de k ($k = 0, 1, 2, \ldots$) vai de zero a infinito. Por exemplo, se $k = 1$, a condição de ângulo torna-se

$$\underline{/s-1} = \mp 540^\circ - 57{,}3^\circ\omega T \qquad \text{(graus)}$$
$$= \mp 3\pi - \omega T \qquad \text{(radianos)}$$

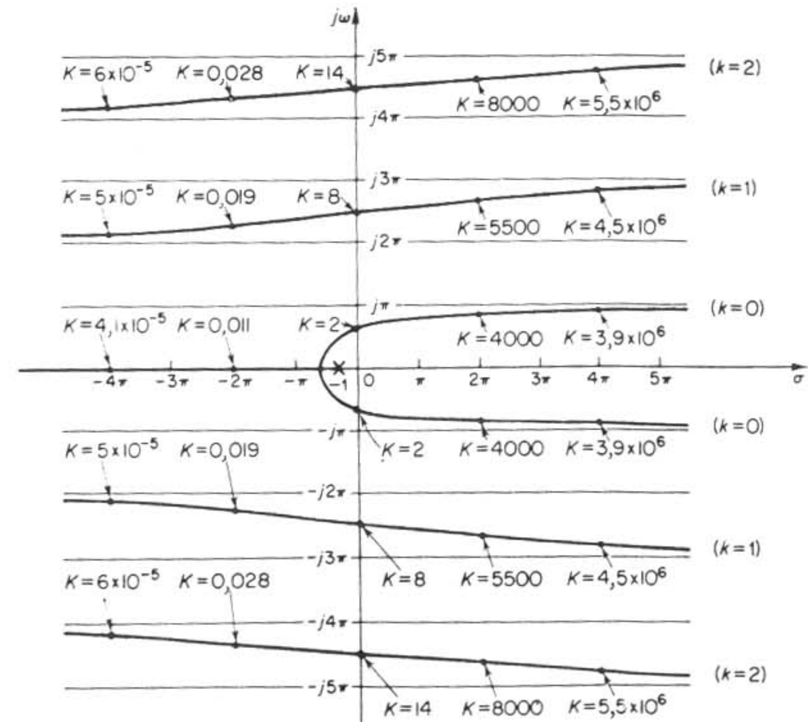


Fig. 8.34 Gráfico do lugar das raízes do sistema indicado na Fig. 8.32 ($T = 1$).

A construção dos lugares das raízes para $k = 1$ é a mesma que para $k = 0$. Um gráfico dos lugares das raízes para $k = 0, 1$ e 2 quando $T = 1$ é indicado na Fig. 8.34.

A condição de módulo estabelece que

$$\left|\frac{Ke^{-Ts}}{s-1}\right| = 1$$

Desde que o módulo de e^{-Ts} é igual àquele de $e^{-T\sigma}$, ou

$$|e^{-Ts}| = e^{-T\sigma}|\cdot|e^{-j\omega T}| = e^{-T\sigma}$$

a condição de módulo resulta

$$|s-1| = Ke^{-T\sigma}$$

Os lugares das raízes indicados na Fig. 8.34 são graduados em termos de K quando $T = 1$. Deste gráfico, é claro que o ramo do lugar das raízes correspondente a $k = 0$ é o dominante; os outros ramos correspondentes a $k = 1, 2, 3, \ldots$ não são tão importantes e podem ser desprezados.

Este exemplo ilustra o fato de que o tempo morto pode causar instabilidade mesmo em sistemas de primeira-ordem porque os lugares das raízes entram no semiplano direito do plano s para valores grandes de K. Portanto, embora o ganho K do sistema de primeira-ordem possa ser ajustado em um valor alto na ausência de tempo morto, ele não pode ser ajustado muito alto se houver tempo morto. (Para o sistema aqui considerado, o valor do ganho K deve ser consideravelmente menor do que 2 para uma operação satisfatória.)

Aproximação de atraso de transporte ou tempo morto. Se o tempo morto T for muito pequeno, então e^{-Ts} poderá ser aproximado por

$$e^{-Ts} \doteq 1 - Ts$$

ou

$$e^{-Ts} \doteq \frac{1}{Ts+1}$$

Estas aproximações são boas se o tempo morto for muito pequeno e, além disso, a função temporal da entrada $f(t)$ para o elemento com tempo morto for uma função suave e contínua. (Isto significa que as derivadas de segunda-ordem e ordens superiores de $f(t)$ são pequenas.)

Efeitos de variações de parâmetros nos pólos de malha-fechada. Em muitos problemas de projeto, os efeitos sobre os pólos de malha-fechada necessitam ser investigados quando são variados outros parâmetros do sistema além do ganho K. Os efeitos destes outros parâmetros podem ser facilmente investigados pelo método do lugar das raízes. Quando dois (ou mais) parâmetros são variados, os lugares das raízes correspondentes são muitas vezes denominados *contornos de raízes*.

Utilizaremos um exemplo para ilustrar a construção dos contornos de raízes quando dois parâmetros são variados, respectivamente, desde zero até infinito.

Considere o sistema indicado na Fig. 8.35. Desejamos investigar o efeito da variação do parâmetro a bem como do ganho K. A função de transferência de

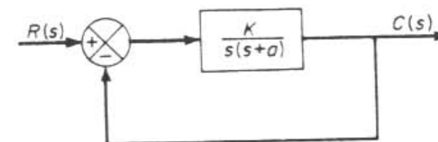


Fig. 8.35 Sistema de controle.

malha-fechada deste sistema é

$$\frac{C(s)}{R(s)} = \frac{K}{s^2 + as + K}$$

A equação característica é

$$s^2 + as + K = 0 \tag{8.19}$$

que pode ser reescrita

$$1 + \frac{as}{s^2 + K} = 0$$

ou

$$\frac{as}{s^2 + K} = -1 \tag{8.20}$$

Na Eq. (8.20), o parâmetro a é escrito como um fator multiplicativo. Para um dado valor de K, o efeito de a nos pólos de malha-fechada podem ser investigados a partir da Eq. (8.20). Os contornos das raízes para este sistema podem ser construídos seguindo-se o procedimento usual para a construção dos lugares das raízes.

Construiremos agora os contornos das raízes conforme K e a variam, respectivamente, desde zero até infinito. Os contornos das raízes começam e terminam nos pólos e zeros de $as/(s^2 + K)$.

Construiremos inicialmente o lugar das raízes quando $a = 0$. Isto pode ser feito facilmente como segue: Substitua $a = 0$ na Eq. (8.19). Então,

$$s^2 + K = 0$$

ou

$$\frac{K}{s^2} = -1 \tag{8.21}$$

Os pólos de malha-aberta, portanto, são pólos duplos na origem. O gráfico do lugar das raízes da Eq. (8.21) é indicado na Fig. 8.36(a).

Para construir os contornos das raízes, vamos inicialmente supor que K é uma constante; por exemplo, $K = 4$. A Eq. (8.20) então resulta

$$\frac{as}{s^2 + 4} = -1 \tag{8.22}$$

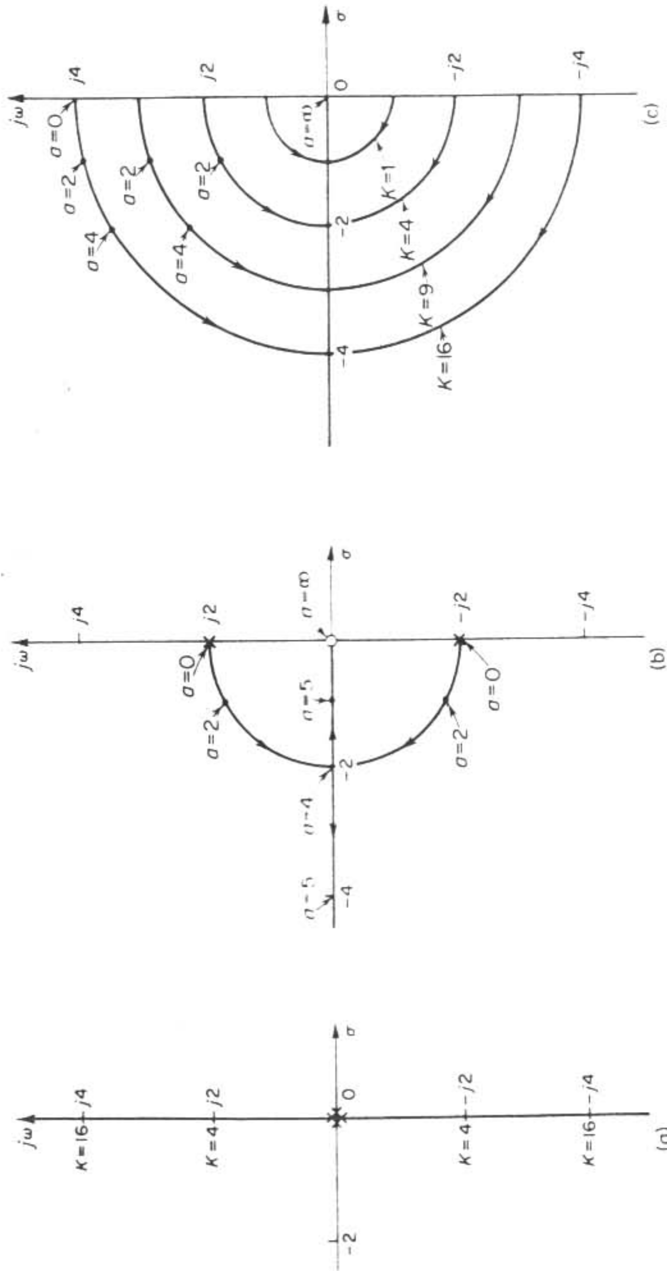


Fig. 8.36 (a) Gráfico do lugar das raízes do sistema indicado na Fig. 8.35. ($a = 0, 0 \leq K \leq \infty$); (b) gráfico do lugar das raízes. ($0 \leq a \leq \infty$, $K = 4$); (c) gráfico do contorno das raízes.

Tabela 8.2 Configurações de pólos e zeros de malha-aberta e lugares das raízes correspondentes

Os pólos de malha-aberta são $s = \pm j2$. O zero de malha-aberta finito está na origem. O gráfico do lugar das raízes correspondente à Eq. (8.22) é indicado na Fig. 8.36(b). Para diferentes valores de K, a Eq. (8.22) fornece os lugares das raízes análogos.

O contorno das raízes, com o diagrama mostrando os lugares das raízes correspondentes a $0 \leq K \leq \infty$, $0 \leq a \leq \infty$, pode ser esboçado como na Fig. 8.36(c). É claro que os contornos das raízes têm início nos pólos e terminam nos zeros da função de transferência $as/(s^2 + K)$. As setas dos contornos das raízes indicam o sentido do aumento no valor de a.

Os contornos das raízes mostram os efeitos das variações dos parâmetros do sistema nos pólos de malha-fechada. Do gráfico do contorno das raízes indicado na Fig. 8.36(c), verificamos que, para $0 < K < \infty$, $0 < a$, os pólos de malha-fechada ficam no semiplano esquerdo do plano s e o sistema é estável.

Note que se o valor de K é fixado, digamos $K = 4$, então os contornos das raízes resultam simplesmente nos lugares das raízes, como mostrado na Fig. 8.36(b).

Ilustramos um método de construção dos contornos das raízes quando o ganho K e o parâmetro a são variados, respectivamente, desde zero até infinito.

Configurações de pólos e zeros típicas e lugares das raízes correspondentes. Concluindo esta seção, forneceremos algumas configurações de pólos e zeros de malha-aberta e seus correspondentes lugares das raízes na Tabela 8.2. A configuração dos lugares das raízes depende apenas da separação relativa dos pólos e zeros de malha-aberta. Se o número de pólos de malha-aberta excede o número de zeros finitos por três ou mais, há um valor do ganho K além do qual os lugares das raízes entram no semiplano direito do plano s, e portanto, o sistema pode tornar-se instável. Um sistema estável deve possuir todos os seus pólos de malha-fechada no semiplano esquerdo do plano s.

Note que uma vez tendo alguma experiência com o método, podemos calcular facilmente as modificações nos lugares das raízes devidas às variações no número e localização dos pólos e zeros de malha-aberta pela visualização dos gráficos do lugar das raízes resultantes de várias configurações de pólos e zeros.

Problemas Ilustrativos e Soluções

Problema A.8.1 Esboce os lugares de ganho constante para o sistema com realimentação unitária, cuja função de transferência do ramo direto é

$$G(s) = \frac{K}{s(s+1)}$$

Solução. Os lugares de ganho constante deste sistema para vários valores de K são dados pela seguinte relação matemática:

$$|G(s)| = \left|\frac{K}{s(s+1)}\right| = 1$$

que pode ser reescrita

$$|s(s+1)| = K \tag{8.23}$$

Vamos substituir $s = \sigma + j\omega$ na Eq. (8.23). Então,

$$\sqrt{\sigma^2 + \omega^2}\sqrt{(\sigma+1)^2 + \omega^2} = K$$

Coloquemos

$$\sigma + \frac{1}{2} = \sigma_1$$

Então, obtemos

$$K^2 = (\sigma^2 + \omega^2)[(\sigma+1)^2 + \omega^2] = \left[\left(\sigma_1 - \frac{1}{2}\right)^2 + \omega^2\right]\left[\left(\sigma_1 + \frac{1}{2}\right)^2 + \omega^2\right]$$

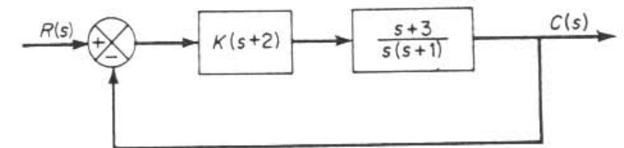
$$= \left(\sigma_1^2 - \frac{1}{4}\right)^2 + 2\omega^2\left(\sigma_1^2 - \frac{1}{4}\right) + \omega^4 + \omega^2$$

ou

$$\left[\left(\sigma_1^2 - \frac{1}{4}\right) + \omega^2\right]^2 = K^2 - \omega^2 \tag{8.24}$$

Os lugares de ganho constante podem ser desenhados pelo uso da Eq. (8.24). Para $K = 1, 2, 5, 10$ e 20, os lugares de ganho constante são indicados na Fig. 8.4.

Problema A.8.2 Esboce o gráfico do lugar das raízes para o sistema indicado na Fig. 8.37(a). (O ganho K é suposto positivo.) Observe que para valores pequenos ou grandes de K o sistema é superamortecido, e para valores médios de K é subamortecido.



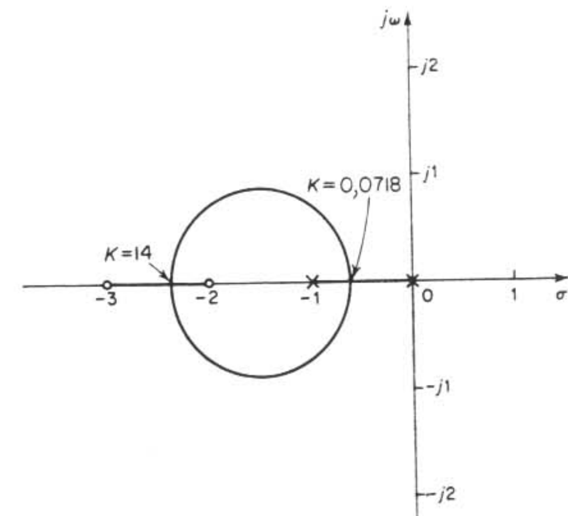


Fig. 8.37 (a) Sistema de controle, (b) gráfico do lugar das raízes.

Solução. O procedimento para construir o gráfico dos lugares das raízes é o seguinte:

1. Coloque os pólos e zeros de malha-aberta no plano complexo. Existem lugares das raízes sobre o eixo real negativo entre 0 e -1 e entre -2 e -3.
2. O número de pólos de malha-aberta e de zeros finitos é o mesmo. Isto significa que não há assíntotas na região complexa do plano s.
3. Determine os pontos de separação de partida e chegada. A equação característica para o sistema é

$$1 + \frac{K(s-2)(s+3)}{s(s+1)} = 0$$

ou

$$K = -\frac{s(s+1)}{(s+2)(s+3)}$$

Os pontos de separação de partida e de chegada são determinados de

$$\frac{dK}{ds} = -\frac{(2s+1)(s+2)(s+3) - s(s+1)(2s+5)}{[(s+2)(s+3)]^2}$$
$$= -\frac{4(s+0{,}634)(s+2{,}366)}{[(s+2)(s+3)]^2}$$
$$= 0$$

como segue:

$$s = -0.634, \qquad s = -2{,}366$$

No ponto $s = -0{,}634$, o valor de K é

$$K = -\frac{(-0{,}634)(0{,}366)}{(1{,}366)(2{,}366)} = 0{,}0718$$

Analogamente, em $s = -2{,}366$,

$$K = -\frac{(-2{,}366)(-1{,}366)}{(-0{,}366)(0{,}634)} = 14$$

Note que os valores de K em $s = -0{,}634$ e $s = -2{,}366$ são positivos e estes pontos são pontos de separação de partida e de chegada verdadeiros. Como o ponto $s = -0{,}634$ ($s = -2.366$) está entre dois pólos (zeros), ele é um ponto de separação de partida (de chegada).
4. Determine um número suficiente de pontos que satisfazem a condição de ângulo. (Pode-se verificar que o lugar das raízes é uma circunferência com centro em $-1{,}5$ que passa pelos pontos de separação de partida e de chegada.) O gráfico do lugar das raízes para este sistema é indicado na Fig. 8.37(b).
5. Usando a condição de módulo, calibre os lugares das raízes em termos de K. Para um dado valor de K, os pólos de malha-fechada, que satisfazem as condições de ângulo e de módulo, podem ser determinados do gráfico do lugar das raízes.

Note que este sistema é estável para qualquer valor positivo de K desde que todos os lugares das raízes permanecem no semiplano esquerdo do plano s.

Pequenos valores de K ($0 < K < 0{,}0718$) correspondem a um sistema superamortecido. Valores médios de K ($0{,}0718 < K < 14$) correspondem a um sistema subamortecido. Finalmente, valores grandes de K ($14 < K$) correspondem a um sistema superamortecido. Com um grande valor de K, pode ser atingido o estado estacionário em um intervalo de tempo muito menor do que com um valor pequeno de K.

O valor de K deve ser ajustado de modo que o desempenho do sistema seja ótimo de acordo com um dado índice de desempenho.

Problema A.8.3 Determine as raízes do seguinte polinômio usando o método do lugar das raízes:

$$3s^4 + 10s^3 + 21s^2 + 24s - 16 = 0 \tag{8.25}$$

Solução. Inicialmente rearranje o polinômio e coloque-o na forma

$$\frac{P(s)}{Q(s)} = -1$$

onde $P(s)$ e $Q(s)$ são polinômios fatorados. Aplicam-se então as regras gerais apresentadas na Seção 8.4 para localizar as raízes do polinômio.

A Eq. (8.25) pode ser rearranjada em um modo conveniente como segue:

$$3s^4 + 10s^3 + 21s^2 = -24s + 16$$

Neste caso o polinômio pode ser reescrito

$$\frac{8(s - \frac{2}{3})}{s^2(s^2 + \frac{10}{3}s + 7)} = -1 \tag{8.26}$$

Esta forma possui dois pólos na origem, dois pólos complexos e um zero no eixo real positivo. Desde que a Eq. (8.26) possui a forma $G(s)H(s) = -1$, o método do lugar das raízes pode ser aplicado para determinar as raízes do polinômio.

A Eq. (8.25) pode ser rearranjada em diferentes maneiras. Por exemplo, pode ser reescrita

$$3s^4 + 10s^3 = -21s^2 - 24s + 16$$

ou

$$\frac{7(s^2 + \frac{8}{7}s - \frac{16}{21})}{s^3(s + \frac{10}{3})} = -1 \tag{8.27}$$

Neste caso, entretanto, o sistema envolve três pólos na origem, um pólo e um zero sobre o eixo real negativo, e um zero no eixo real positivo. A quantidade de trabalho gráfico necess[illegible] ria para esboçar o gráfico do lugar das raízes da Eq. (8.27) é quase a mesma daque[illegible] correspondente à Eq. (8.26).

Note que se dois gráficos dos lugares das raízes, um correspondente à Eq. (8.26) e out[illegible] correspondente à Eq. (8.27), são esboçados no mesmo diagrama, as intersecções dos do[illegible] fornecem as raízes dos polinômios. (Se for utilizada a condição de módulo, apenas um gráfi[illegible] do lugar das raízes necessita ser esboçado.)

Neste problema, esboçaremos apenas o gráfico do lugar das raízes baseado na Eq. (8.[illegible] e utilizaremos a condição do módulo para determinar as raízes do polinômio. A Eq. (8.[illegible] pode ser reescrita

$$\frac{8(s - \frac{2}{3})}{s^2(s + 1{,}67 + j2{,}06)(s + 1{,}67 - j2{,}06)} = -1 \qquad (8.[illegible]$$

Para determinar os lugares das raízes, substitua a constante 8 no numerador da Eq. (8.28) por K e escreva

$$\frac{K(s - \frac{2}{3})}{s^2(s + 1{,}67 + j2{,}06)(s + 1{,}67 - j2{,}06)} = -1$$

Para esboçar o gráfico do lugar das raízes, siga este procedimento:
1. Coloque os pólos e o zero no plano complexo. Existem lugares das raízes sobre o eixo real entre 2/3 e 0 e entre 0 e $-\infty$.
2. Determine as assíntotas dos lugares das raízes. Há três assíntotas que fazem ângulos de

$$\frac{\pm 180°(2k + 1)}{4 - 1} = 60°, -60°, 180°$$

com o eixo real positivo. Referindo-se à Eq. (8.13) a abscissa da intersecção das assíntotas é dada por

$$-\sigma_a = -\frac{(0 + 0 + \frac{5}{3} + j2{,}06 + \frac{5}{3} - j2{,}06) + \frac{2}{3}}{4 - 1} = -\frac{4}{3}$$

3. Usando o critério de estabilidade de Routh, determine o valor de K para o qual o lugar das raízes cruza o eixo imaginário. A equação característica é

$$s^2(s^2 + \tfrac{10}{3}s + 7) = -K(s - \tfrac{2}{3})$$

ou

$$s^4 + \tfrac{10}{3}s^3 + 7s^2 + Ks - \tfrac{2}{3}K = 0$$

O arranjo de Routh resulta

$$\begin{array}{llll}
s^4 & 1 & 7 & -\frac{2}{3}K \\
s^3 & \frac{10}{3} & K & 0 \\
s^2 & 7 - \frac{3}{10}K & -\frac{2}{3}K & \\
s^1 & \dfrac{-\frac{3}{10}K^2 + \frac{83}{9}K}{7 - \frac{3}{10}K} & 0 & \\
s^0 & -\frac{2}{3}K & &
\end{array}$$

Os valores de K que tornam nulo o termo s^1 na primeira coluna são $K = 30{,}7$ e $K = 0$. Os pontos de cruzamento no eixo imaginário podem ser determinados resolvendo-se a equação auxiliar obtida da linha s^2, ou

$$(7 - \tfrac{3}{10}K)s^2 - \tfrac{2}{3}K = 0$$

onde $K = 30{,}7$. O resultado é

$$s = \pm j3{,}04$$

Os pontos de cruzamento sobre o eixo imaginário são, portanto, $s = \pm j3{,}04$.
4. Verifique se há pontos de separação de partida ou chegada. Do Passo 3 sabemos que há dois pontos de cruzamento com o eixo $j\omega$. Portanto, nesta configuração

particular de pólos e zeros, não pode existir qualquer ponto de separação de partida ou chegada.
5. Determine os ângulos de partida dos lugares das raízes dos pólos complexos. No pólo $s = -1.67 + j\,2.06$, o ângulo de partida θ é determinado de

$$110° - 106° - 106° - 90° - \theta = \pm 180°(2k + 1)$$

como segue:

$$\theta = -12°$$

(O ângulo de partida do pólo $s = -1.67 - j2.06$ é 12°.)
6. Na vizinhança do eixo $j\omega$ e da origem, localize um número suficiente de pontos que satisfaça a condição do ângulo.

Baseado na informação obtida até aqui, o gráfico do lugar das raízes para este sistema pode ser esboçado conforme a Fig. 8.38.
7. Use a condição do módulo

$$K = \left| \frac{s^2(s + 1{,}67 + j2{,}06)(s + 1{,}67 - j2{,}06)}{s - \frac{2}{3}} \right|$$

para determinar os pontos sobre o lugar das raízes nos quais $K = 8$. Utilizando um procedimento de tentativa-e-erro, obtemos

$$s = -0{,}79 \pm j2{,}16$$

Pode ser utilizado um procedimento de tentativa-e-erro para localizar as duas raízes

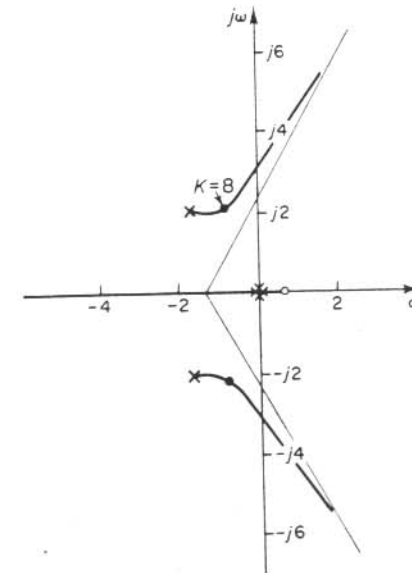


Fig. 8.38 Gráfico do lugar das raízes.

restantes. Entretanto, pode ser mais simples fatorar as raízes conhecidas do polinômio dado.

$$3s^4 + 10s^3 + 21s^2 + 24s - 16$$
$$= (s + 0{,}79 + j2{,}16)(s + 0{,}79 - j2{,}16)(3s^2 + 5{,}28s - 3{,}06)$$
$$= 3(s + 0{,}79 + j2{,}16)(s + 0{,}79 - j2{,}16)(s + 2{,}22)(s - 0{,}46)$$

Portanto, as raízes do polinômio dado são

$$s_1 = -0{,}79 - j2{,}16, \quad s_2 = -0{,}79 + j2{,}16, \quad s_3 = -2{,}22, \quad s_4 = 0{,}46$$

Problema A.8.4 Uma forma simplificada da função de transferência de malha-aberta de um avião com piloto automático no modo longitudinal é

$$G(s)H(s) = \frac{K(s + a)}{s(s - b)(s^2 + 2\zeta\omega_n s + \omega_n^2)}, \quad a > 0, \quad b > 0$$

Um sistema deste tipo envolvendo um pólo de malha-aberta no semiplano direito do plano s pode ser condicionalmente estável. Esboce o gráfico do lugar das raízes quando $a = b = 1$, $\zeta = 0{,}5$ e $\omega_n = 4$. Determinar o intervalo do ganho K para estabilidade.

Solução. A função de transferência de malha-aberta para o sistema é

$$G(s)H(s) = \frac{K(s + 1)}{s(s - 1)(s^2 + 4s + 16)}$$

Esboçar o gráfico do lugar das raízes, seguindo este procedimento:
1. Indique no plano complexo os pólos e zeros de malha-aberta. Existem lugares das raízes sobre o eixo real entre 1 e 0 e entre -1 e $-\infty$.
2. Determine as assíntotas dos lugares das raízes. Há três assíntotas cujos ângulos podem ser determinados como

$$\text{Ângulos das assíntotas} = \frac{180°(2k + 1)}{4 - 1} = 60°, -60°, 180°$$

A abscissa da intersecção das assíntotas com o eixo real é

$$-\sigma_a = -\frac{(0 - 1 + 2 + j2\sqrt{3} + 2 - j2\sqrt{3}) - 1}{4 - 1} = -\frac{2}{3}$$

3. Determine os pontos de separação de partida e de chegada. Desde que a equação característica é

$$1 + \frac{K(s + 1)}{s(s - 1)(s^2 + 4s + 16)} = 0$$

obtemos

$$K = -\frac{s(s - 1)(s^2 + 4s + 16)}{s + 1}$$

Diferenciando K em relação a s, obtemos

$$\frac{dK}{ds} = -\frac{3s^4 + 10s^3 + 21s^2 + 24s - 16}{(s + 1)^2}$$

No Problema A.8.3, verificamos que

$$3s^4 + 10s^3 + 21s^2 + 24s - 16$$
$$= 3(s + 0{,}79 + j2{,}16)(s + 0{,}79 - j2{,}16)(s + 2{,}22)(s - 0{,}46)$$

Portanto, os pontos de separação de partida e de chegada são $s = 0{,}46$ e $s = -2{,}22$, respectivamente. (Os pontos $s = -0{,}79 \pm j2{,}16$ não satisfazem a condição do ângulo.)
4. Utilizando o critério de estabilidade de Routh, determine o valor de K para o qual os lugares das raízes cruzam o eixo imaginário. Desde que a equação característica é

$$s^4 + 3s^3 + 12s^2 + (K - 16)s + K = 0$$

o arranjo de Routh resulta

s^4	1	12	K
s^3	3	$K - 16$	0
s^2	$\frac{52 - K}{3}$	K	0
s^1	$\frac{-K^2 + 59K - 832}{3(52 - K)}$	0	
s^0	K		

Os valores de K que anulam o termo s^1 na primeira coluna são $K = 35{,}7$ e $K = 23{,}3$.

Os pontos de cruzamento do eixo imaginário podem ser determinados resolvendo-se a equação auxiliar obtida da linha s^2, isto é, resolvendo a seguinte equação em relação a s:

$$\frac{52 - K}{3}s^2 + K = 0$$

Os resultados são

$$s = \pm j2{,}56 \quad \text{para } K = 35{,}7$$

$$s = \pm j1{,}56 \quad \text{para } K = 23{,}3$$

Os pontos de cruzamento sobre o eixo imaginário são portanto $s = \pm j2{,}56$ e $s = \pm j1{,}56$.
5. Determine os ângulos de partida dos lugares das raízes dos pólos complexos. Para o pólo de malha-aberta em $s = -2 + j2\sqrt{3}$, o ângulo de partida θ é

$$106° - 120° - 130{,}5° - 90° - \theta = \pm 180°(2k + 1)$$

ou

$$\theta = -54{,}5°$$

(O ângulo de partida do pólo de malha-aberta em $s = -2 - j2\sqrt{3}$ é $54{,}5°$.)
6. Escolha um ponto de teste na vizinhança do eixo $j\omega$ e da origem e aplique a condição do ângulo. Se o ponto de teste não satisfizer a condição do ângulo, selecione um outro

ponto de teste até satisfazê-la. Continue o mesmo processo e localize um número suficiente de pontos que satisfaça a condição do ângulo.

A Fig. 8.39 mostra o gráfico do lugar das raízes para este sistema. Do Passo 4 o sistema é estável para $23,3 < K < 35,7$. Caso contrário, é instável.

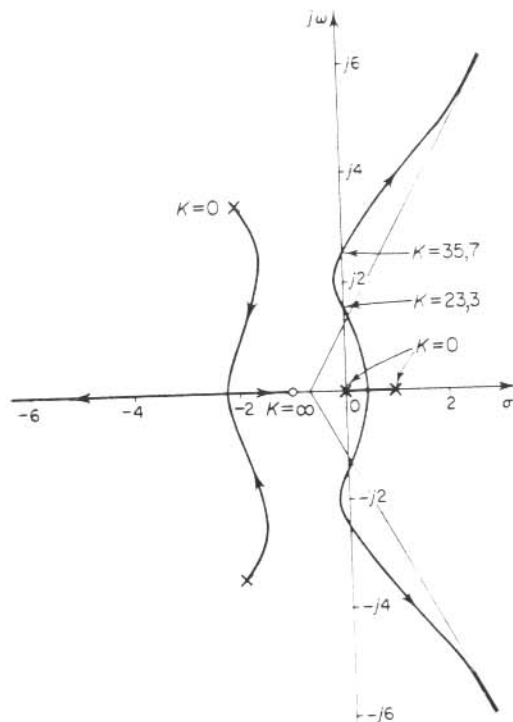


Fig. 8.39 Gráfico do lugar das raízes.

Problema A.8.5 Considere o sistema com atraso de transporte indicado na Fig. 8.40. Esboce o gráfico do lugar das raízes e determine os dois pares de pólos de malha-fechada mais próximos do eixo $j\omega$.

Utilizando apenas os pólos de malha-fechada dominantes, obtenha a resposta ao degrau unitário e esboce a curva da resposta.

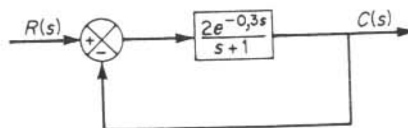


Fig. 8.40 Sistema de controle com atraso de transporte.

Solução. A equação característica é

$$\frac{2e^{-0,3s}}{s+1} + 1 = 0$$

que é equivalente às seguintes condições de ângulo e módulo:

$$\angle \frac{2e^{-0,3s}}{s+1} = \pm 180°(2k+1)$$

$$\left|\frac{2e^{-0,3s}}{s+1}\right| = 1$$

A condição do ângulo se reduz a

$$\angle s+1 = \mp\pi(2k+1) - 0,3\omega \quad \text{(radianos)}$$

Para $k = 0$,

$$\angle s+1 = \mp\pi - 0,3\omega \quad \text{(radianos)}$$
$$= \mp 180° - 17,2°\omega \quad \text{(graus)}$$

Para $k = 1$,

$$\angle s+1 = \mp 3\pi - 0,3\omega \quad \text{(radianos)}$$
$$= \mp 540° - 17,2°\omega \quad \text{(graus)}$$

O gráfico do lugar das raízes para este sistema é indicado na Fig. 8.41.

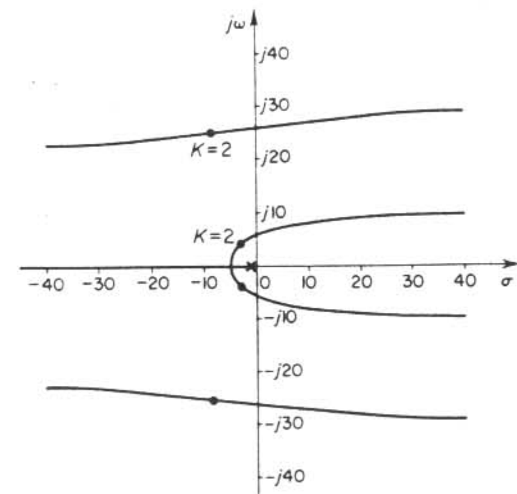


Fig. 8.41 Gráfico do lugar das raízes para o sistema indicado na Fig. 8.40.

Vamos substituir $s = \sigma + j\omega$ na condição do módulo e substituir 2 por K. Obtemos então

$$\frac{\sqrt{(1+\sigma)^2+\omega^2}}{e^{-0,3\sigma}} = K$$

Calculando K em diferentes pontos dos lugares das raízes podem-se determinar os pontos correspondentes a $K = 2$. Estes pontos são os pólos de malha-fechada. O par dominante de pólos de malha-fechada é

$$s = -2,5 \pm j3,9$$

O par de pólos de malha-fechada seguinte é

$$s = -8,6 \pm j25,1$$

Considerando apenas o par de pólos de malha-fechada dominantes, a função de transferência de malha-fechada pode ser aproximada como segue: Verifique que

$$\frac{C(s)}{R(s)} = \frac{2e^{-0,3s}}{1+s+2e^{-0,3s}}$$

$$= \frac{2e^{-0,3s}}{1+s+2\left(1-0,3s+\frac{0,09s^2}{2}+\cdots\right)}$$

$$= \frac{2e^{-0,3s}}{3+0,4s+0,09s^2+\cdots}$$

e

$$(s+2,5+j3,9)(s+2,5-j3,9) = s^2+5s+21,46$$

Podemos aproximar $C(s)/R(s)$ por

$$\frac{C(s)}{R(s)} = \frac{\frac{2}{3}(21,46)e^{-0,3s}}{s^2+5s+21,46}$$

ou

$$\frac{C(s)}{R(s)} = \frac{14,31e^{-0,3s}}{(s+2,5)^2+3,9^2}$$

Para uma entrada em degrau unitário.

$$C(s) = \frac{14,31e^{-0,3s}}{[(s+2,5)^2+3,9^2]s}$$

Note que

$$\frac{14,31}{[(s+2,5)^2+3,9^2]s} = \frac{\frac{2}{3}}{s} + \frac{-\frac{2}{3}s-\frac{10}{3}}{(s-2,5)^2+3,9^2}$$

Portanto,

$$C(s) = \left(\frac{\frac{2}{3}}{s}\right)e^{-0,3s} + \left[\frac{-\frac{2}{3}s-\frac{10}{3}}{(s+2,5)^2+3,9^2}\right]e^{-0,3s}$$

A transformada de Laplace inversa de $C(s)$ resulta

$$c(t) = \tfrac{2}{3}[1 - e^{-2,5(t-0,3)}\cos 3,9(t-0,3) - 0,64e^{-2,5(t-0,3)}\,\text{sen}\,3,9(t-0,3)]1(t-0,3)$$

onde $1(t - 0.3)$ é a função degrau unitário ocorrendo em $t = 0.3$.

A Fig. 8.42 mostra a curva de resposta aproximada obtida, junto com a curva de resposta ao degrau unitário exata obtida por cálculo numérico. Note que neste sistema pode-se obter uma aproximação muito boa usando os pólos de malha-fechada dominantes.

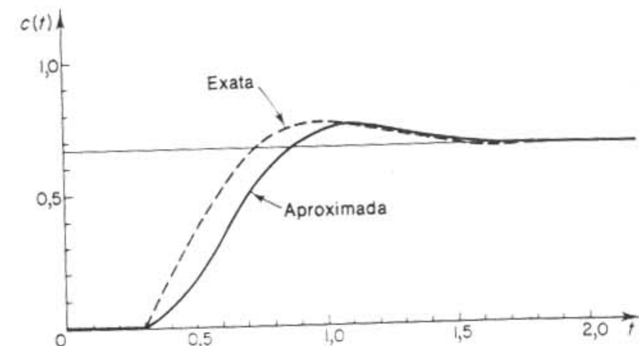


Fig. 8.42 Curvas de resposta ao degrau unitário para o sistema indicado na Fig. 8.40.

Problema A.8.6* Em um sistema de controle complexo, pode haver uma realimentação positiva em um laço interno conforme indicado na Fig. 8.43. Este laço normalmente é estabilizado pelo laço mais externo.

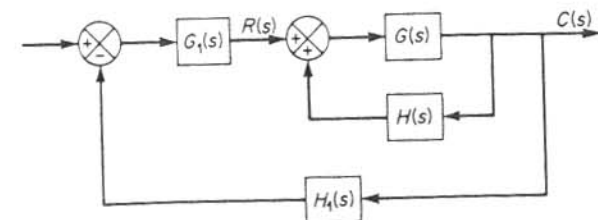


Fig. 8.43 Sistema de controle.

*Referência W-4.

Neste problema nos preocuparemos apenas com o laço interno de realimentação positiva. A função de transferência de malha-fechada do laço interno é

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 - G(s)H(s)}$$

A equação característica é

$$1 - G(s)H(s) = 0 \tag{8.29}$$

Esta equação pode ser resolvida de maneira análoga ao desenvolvimento do método do lugar das raízes nas Seções 8.3 e 8.4. A condição do ângulo, entretanto, deve ser alterada.

A Eq. (8.29) pode ser reescrita:

$$G(s)H(s) = 1$$

que é equivalente às duas equações seguintes:

$$\underline{/G(s)H(s)} = 0^\circ \pm k360^\circ \qquad (k = 0, 1, 2, \ldots)$$

$$|G(s)H(s)| = 1$$

A soma total de todos os ângulos dos pólos e zeros de malha-aberta deve ser igual a $0^\circ \pm k$ 360°. Portanto, o lugar das raízes segue um lugar de 0° em oposição ao lugar de 180° anteriormente considerado. A condição do módulo permanece inalterada.

Esboce o gráfico do lugar das raízes para o sistema de realimentação positiva com as seguintes funções de transferência:

$$G(s) = \frac{K(s+2)}{(s+3)(s^2+2s+2)},$$

$$H(s) = 1$$

O ganho K é suposto positivo.

Solução. As regras gerais para construção dos lugares das raízes fornecidas na Seção 8.4 devem ser modificadas da seguinte maneira:

A regra 3 é modificada como segue: Se o número total de pólos reais e zeros reais à direita de um ponto de teste sobre o eixo real for par, então este ponto pertence ao lugar das raízes.

A regra 4 é modificada como segue:

$$\text{Ângulos dos assíntotas} = \frac{\pm k360^\circ}{n - m}$$

onde

n = número de pólos finitos de $G(s)H(s)$
m = número de zeros finitos de $G(s)H(s)$

A regra 6 é modificada como segue: Quando calcular o ângulo de partida (ou o ângulo de chegada) de um pólo complexo (ou um zero complexo) de malha-aberta, subtrair de 0° a soma de todos os ângulos dos números complexos obtidos de todos os outros pólos e zeros para o pólo complexo (ou zero complexo) em questão, incluindo-se os sinais apropriados.

As outras regras para construção do gráfico do lugar das raízes permanecem inalteradas. Aplicaremos agora as regras modificadas para construir o gráfico do lugar das raízes. A

função de transferência de malha-fechada para o sistema com realimentação positiva é dada por

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 - G(s)H(s)} = \frac{K(s+2)}{(s+3)(s^2+2s+2) - K(s+2)}$$

1. Indique os pólos de malha-aberta ($s = -1 + j$, $s = -1 - j$, $s = -3$) e o zero de malha-aberta ($s = -2$) no plano complexo. Conforme K é aumentado desde 0 até ∞, os pólos de malha-fechada iniciam-se nos pólos de malha-aberta e terminam nos zeros de malha-aberta (finitos ou infinitos), como no caso dos sistemas com realimentação negativa.
2. Determine os lugares das raízes sobre o eixo real. Existem lugares das raízes sobre o eixo real entre -2 e $+\infty$ e entre -3 e $-\infty$.
3. Determine as assíntotas dos lugares das raízes. Para o presente sistema,

$$\text{Ângulo da assíntota} = \frac{\pm k360^\circ}{3 - 1} = \pm 180^\circ$$

Isto significa simplesmente que os ramos dos lugares das raízes estão sobre o eixo real.
4. Determine os pontos de separação de partida e de chegada. Desde que a equação característica é

$$(s+3)(s^2+2s+2) - K(s+2) = 0$$

obtemos

$$K = \frac{(s+3)(s^2+2s+2)}{s+2}$$

Diferenciando K em relação a s, obtemos

$$\frac{dK}{ds} = \frac{2s^3 + 11s^2 + 20s + 10}{(s+2)^2}$$

Note que

$$\begin{aligned} 2s^3 + 11s^2 + 20s + 10 &= 2(s+0{,}8)(s^2 + 4{,}7s + 6{,}24) \\ &= 2(s+0{,}8)(s + 2{,}35 + j0{,}77)(s + 2{,}35 - j0{,}77) \end{aligned}$$

Os pontos $s = -2{,}35 \pm j0{,}77$ não satisfazem a condição do ângulo. No ponto $s = -0{,}8$, o valor de K resulta positivo. O ponto de separação de chegada é portanto $s = -0{,}8$.
5. Determine o ângulo de partida do lugar das raízes de um pólo complexo. Para o pólo complexo em $s = -1 + j$, o ângulo de partida ϕ é

$$45^\circ - 27^\circ - 90^\circ - \phi = 0^\circ$$

ou

$$\phi = -72^\circ$$

(O ângulo de partida do pólo complexo em $s = -1 - j$ é 72°.)
6. Escolha um ponto de teste na vizinhança do eixo $j\omega$ e da origem e aplique a condição do ângulo. Localize um número suficiente de pontos que satisfazem a condição do ângulo.

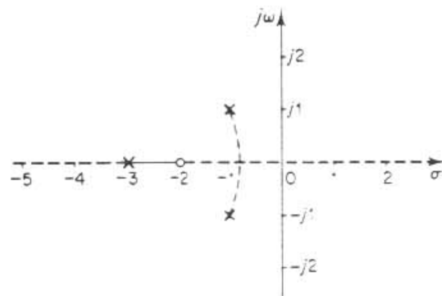


Fig. 8.44 Gráfico do lugar das raízes para o sistema com realimentação positiva com $G(s) = K(s + 2)/[(s + 3)(s^2 + 2s + 2)]$, $H(s) = 1$.

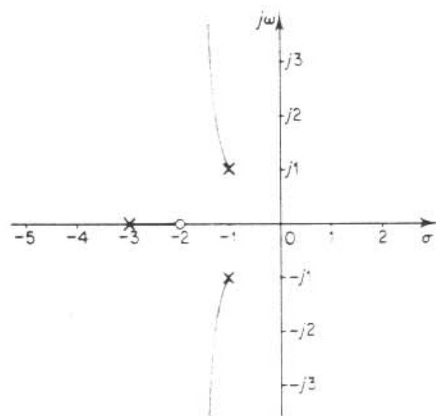


Fig. 8.45 Gráfico do lugar das raízes para o sistema com realimentação negativa com $G(s) = K(s + 2)/[(s + 3)(s^2 + 2s + 2)]$, $H(s) = 1$.

A Fig. 8.44 mostra o gráfico do lugar das raízes para o sistema com realimentação positiva dado. Os lugares das raízes são indicados por retas e curvas tracejadas.

Note que, se

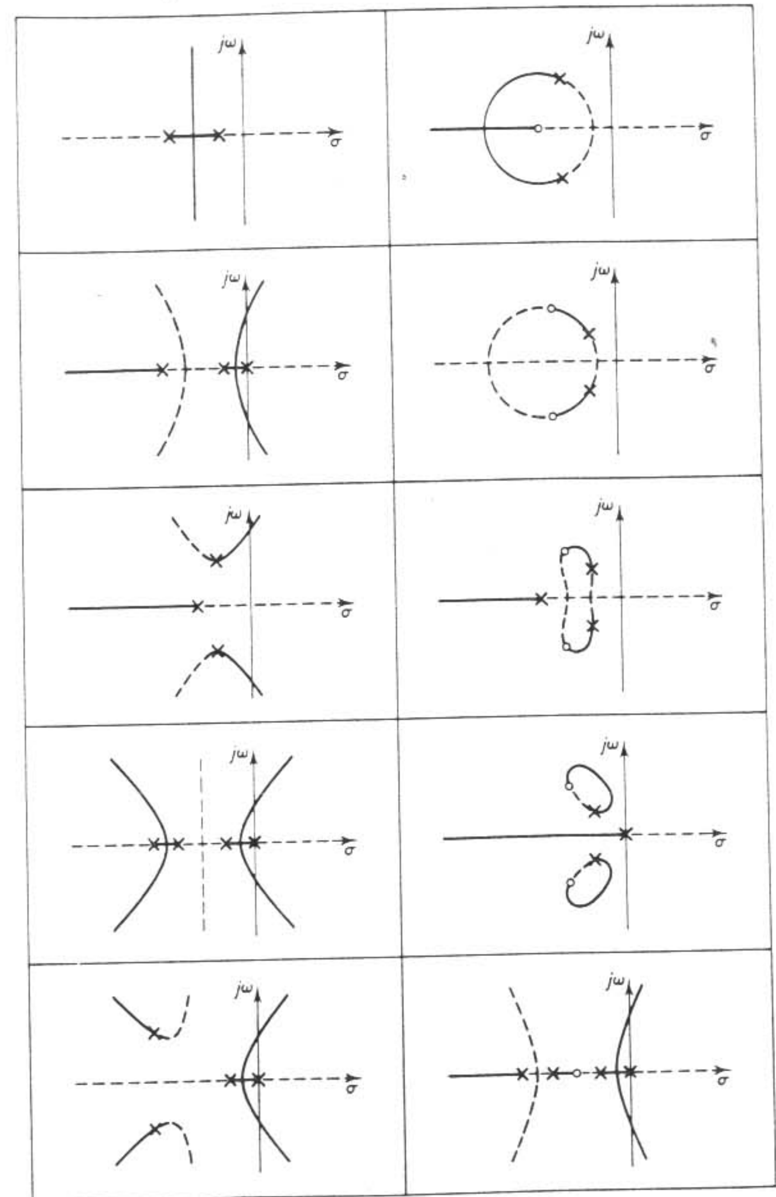
$$K > \left. \frac{(s + 3)(s^2 + 2s + 2)}{s + 2} \right|_{s=0} = 3$$

uma raiz real entra no semiplano direito do plano s. Portanto, para valores de K maiores do que 3, o sistema torna-se instável. (Para $K > 3$, o sistema deve ser estabilizado com um laço externo.)

Para comparar o gráfico do lugar das raízes com aquele do sistema de realimentação negativa correspondente, mostramos na Fig. 8.45 o gráfico do lugar das raízes do sistema com realimentação cuja função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{K(s + 2)}{(s + 3)(s^2 + 2s + 2) + K(s + 2)}$$

Tabela 8.3 Gráficos dos lugares das raízes de sistemas com realimentação negativa e com realimentação positiva



Retas e curvas cheias correspondem a sistemas com realimentação negativa; retas e curvas tracejadas correspondem a sistemas com realimentação positiva.

A Tabela 8.3 mostra vários gráficos de lugar das raízes de sistemas com realimentação negativa e sistemas com realimentação positiva. As funções de transferência de malha-fechada são dadas por

$$\frac{C}{R} = \frac{G}{1 + GH} \quad \text{para sistemas com realimentação negativa}$$

$$\frac{C}{R} = \frac{G}{1 - GH} \quad \text{para sistemas com realimentação positiva}$$

onde GH é a função de transferência de malha-aberta. Na Tabela 8.3, os lugares das raízes para sistemas com realimentação negativa são desenhados com retas e curvas cheias, e aqueles para sistemas com realimentação positiva são desenhados com retas e curvas tracejadas.

Problema A.8.7 Considere o sistema mostrado na Fig. 8.46, que tem uma função de transferência do ramo direto instável. Esboce o gráfico do lugar das raízes e localize os pólos de malha-fechada. Mostre que embora os pólos de malha-fechada estejam sobre o eixo real negativo e o sistema não seja oscilatório, a curva de resposta ao degrau unitário possui sobrelevação.

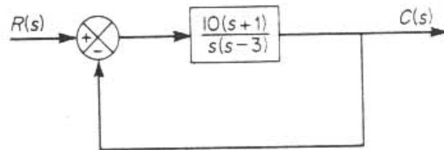


Fig. 8.46 Sistema de controle.

Solução. O gráfico do lugar das raízes para este sistema é indicado na Fig. 8.47. Os pólos de malha-fechada estão localizados em $s = -2$ e $s = -5$.

A função de transferência de malha-fechada resulta

$$\frac{C(s)}{R(s)} = \frac{10(s+1)}{s^2 + 7s + 10}$$

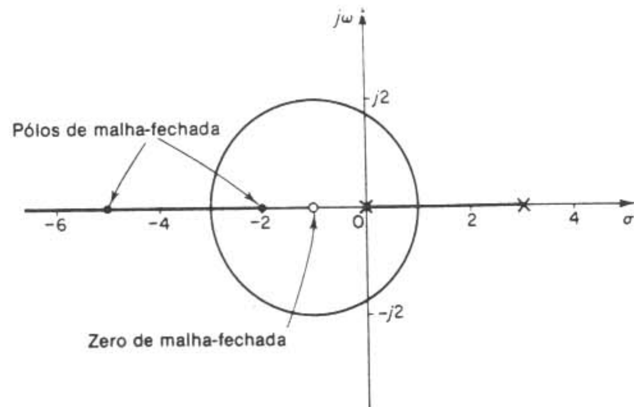


Fig. 8.47 Gráfico do lugar das raízes para o sistema indicado na Fig. 8.46.

A resposta ao degrau unitário deste sistema é

$$C(s) = \frac{10(s+1)}{s(s+2)(s+5)}$$

A transformada de Laplace inversa de $C(s)$ resulta

$$c(t) = 1 + 1{,}666e^{-2t} - 2{,}666e^{-5t} \qquad (t \geq 0)$$

A curva de resposta ao degrau unitário é indicada na Fig. 8.48. Embora o sistema não seja oscilatório, a curva de resposta ao degrau unitário exibe sobrelevação. (Isto é devido à presença de um zero em $s = -1$.)

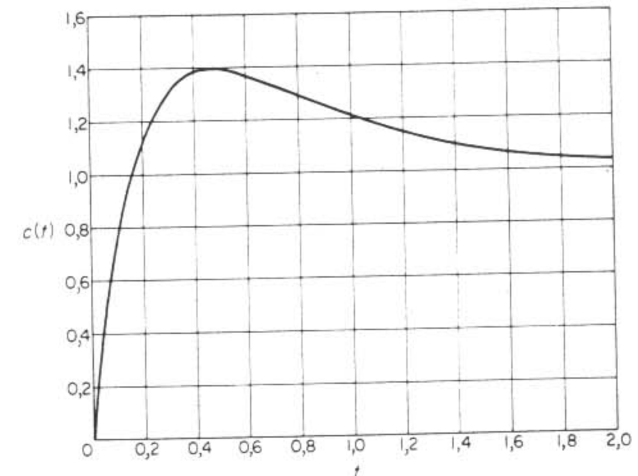


Fig. 8.48 Curva de resposta ao degrau unitário para o sistema indicado na Fig. 8.46.

Problema A.8.8 Considere um sistema com

$$G(s) = \frac{s+3}{s(s+2)}, \qquad H(s) = 1$$

Mostre que uma parte do lugar das raízes é circular.

Solução. Sobre o lugar das raízes, a condição do ângulo resulta

$$\underline{/G(s)} = \underline{/s+3} - \underline{/s} - \underline{/s+2} = 180°$$

Se $s = \sigma + j\omega$ é substituído nesta última equação, obtemos

$$\underline{/\sigma + j\omega + 3} - \underline{/\sigma + j\omega} - \underline{/\sigma + j\omega + 2} = 180°$$

que pode ser reescrita

$$\tan^{-1}\frac{\omega}{\sigma+3}-\tan^{-1}\frac{\omega}{\sigma}=180^\circ+\tan^{-1}\frac{\omega}{\sigma+2}$$

Considerando as tangentes de ambos os membros desta última equação e usando a relação

$$\tan(x\pm y)=\frac{\tan x\pm\tan y}{1\pm\tan x\tan y}$$

obtemos

$$\tan\left(\tan^{-1}\frac{\omega}{\sigma+3}-\tan^{-1}\frac{\omega}{\sigma}\right)=\frac{\dfrac{\omega}{\sigma+3}-\dfrac{\omega}{\sigma}}{1+\dfrac{\omega}{\sigma+3}\dfrac{\omega}{\sigma}}=\frac{-3\omega}{\sigma(\sigma+3)+\omega^2}$$

$$\tan\left(180^\circ+\tan^{-1}\frac{\omega}{\sigma+2}\right)=\frac{0+\dfrac{\omega}{\sigma+2}}{1-0\times\dfrac{\omega}{\sigma+2}}=\frac{\omega}{\sigma+2}$$

Portanto,

$$\frac{-3\omega}{\sigma(\sigma+3)+\omega^2}=\frac{\omega}{\sigma+2}$$

ou

$$(\sigma+3)^2+\omega^2=(\sqrt{3})^2$$

Esta última equação representa uma circunferência com centro em $\sigma=-3$, $\omega=0$ e com raio $\sqrt{3}$. Note que o centro está no zero da função de transferência de malha-aberta. A Fig. 8.49 mostra o gráfico do lugar das raízes para o sistema.

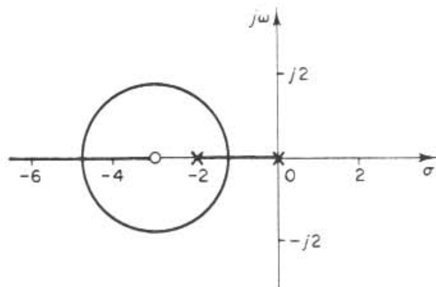


Fig. 8.49 Gráfico do lugar das raízes para o sistema com $G(s)=(s+3)/[s(s+2)]$, $H(s)=1$.

PROBLEMAS

Problema B.8.1 Esboce os gráficos dos lugares das raízes para as configurações de pólos e zeros de malha-aberta indicadas na Fig. 8.50.

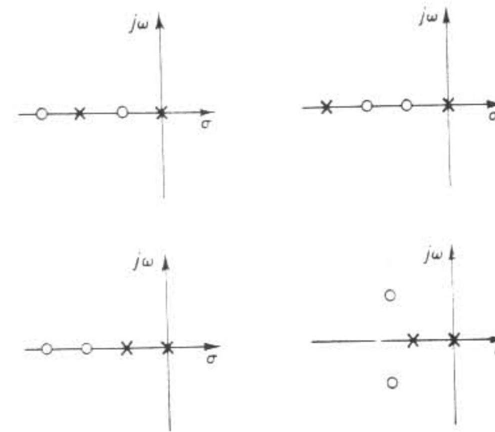


Fig. 8.50 Pólos e zeros de sistemas.

Problema B.8.2 Esboce os gráficos dos lugares das raízes para as configurações de pólos e zeros de malha-aberta indicadas na Fig. 8.51.

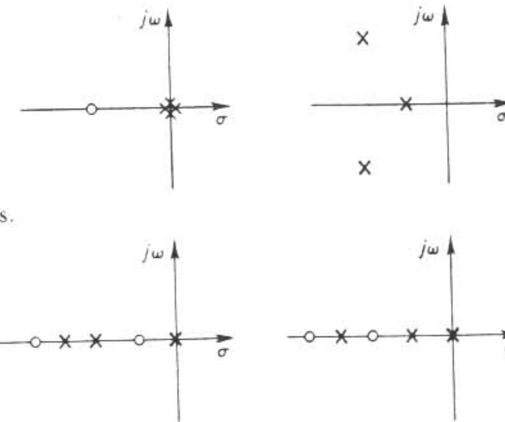


Fig. 8.51 Pólos e zeros de sistemas.

Problema B.8.3 Considere o sistema indicado na Fig. 8.52. Investigue o efeito do aumento do valor de K_h nos lugares das raízes. Esboce as curvas típicas de resposta ao degrau unitário para os casos $K_h=0$, $0<K_h<1$ e $K_h>1$. Esboce então o gráfico do lugar das raízes quando $K_h=0{,}5$. Localize os pólos de malha-fechada nos lugares das raízes quando $K=10$.

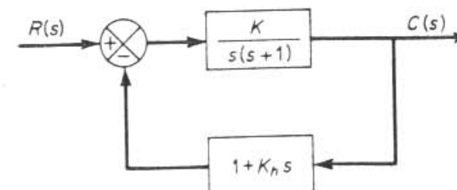


Fig. 8.52 Sistema de controle.

Problema B.8.4 Mostre que os lugares das raízes para um sistema de controle com

$$G(s) = \frac{K(s^2 + 6s + 10)}{s^2 + 2s + 10},$$

$$H(s) = 1$$

são arcos de circunferências centradas na origem e com raio igual a $\sqrt{10}$.

Problema B.8.5 Um sistema de controle com

$$G(s) = \frac{K}{s^2(s + 1)},$$

$$H(s) = 1$$

é instável para todos os valores positivos do ganho K.

Esboce o gráfico do lugar das raízes deste sistema. Usando e te gráfico, mostre que este sistema pode ser estabilizado adicionando-se um zero sobre o eixo real negativo ou modificando-se $G(s)$ para $G_1(s)$, onde

$$G_1(s) = \frac{K(s + a)}{s^2(s + 1)} \qquad (0 \leq a < 1)$$

Problema B.8.6 Esboce o gráfico do lugar das raízes de um sistema com

$$G(s) = \frac{K}{(s^2 + 2s + 2)(s^2 + 2s + 5)},$$

$$H(s) = 1$$

Determine os pontos exatos onde os lugares das raízes cruzam o eixo $j\omega$.

Problema B.8.7 Esboce o gráfico do lugar das raízes para um sistema com

$$G(s) = \frac{K}{s(s + 0{,}5)(s^2 + 0{,}6s + 10)},$$

$$H(s) = 1$$

Problema B.8.8 Esboce o gráfico do lugar das raízes para o sistema indicado na Fig. 8.53 e demonstre que ele se torna instável para valores grandes de K.

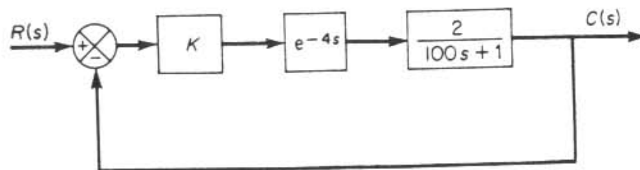


Fig. 8.53 Sistema de controle.

Problema B.8.9 Esboce os contornos das raízes para o sistema indicado na Fig. 8.54 quando o ganho K e o parâmetro a variam, respectivamente, desde zero até infinito.

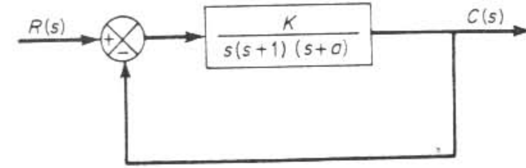


Fig. 8.54 Sistema de controle.

Problema B.8.10 Conside e o sistema indicado na Fig. 8.55. Determine os valores do ganho K e o coeficiente de realime .tação de velocidade K_h de modo que os pólos de malha-fechada sejam $s = -1 \pm j\sqrt{3}$. Posteriormente, usando o valor de K_h determinado, esboce o gráfico do lugar das raízes.

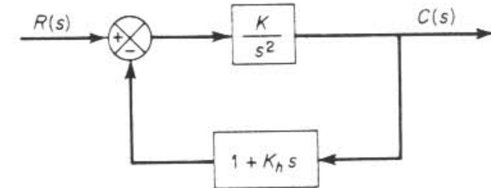


Fig. 8.55 Sistema de controle.

9

Métodos de Resposta em Freqüência

9.1 INTRODUÇÃO

Resposta em freqüência. O termo "resposta em freqüência" significa a resposta em regime estacionário de um sistema com entrada senoidal. Nos métodos de resposta em freqüência, o método mais convencional disponível aos engenheiros de controle para análise e projeto de sistemas de controle, variamos a freqüência do sinal de entrada em um certo intervalo e estudamos a resposta em freqüência resultante.

O critério de estabilidade de Nyquist, a ser apresentado na Seção 9.5, possibilita investigar tanto a estabilidade absoluta como a estabilidade relativa de sistemas de malha-fechada lineares a partir do conhecimento de suas características de resposta em freqüência de malha-aberta. Usando este critério de estabilidade, não precisamos determinar as raízes da equação característica. Esta é uma vantagem da abordagem de resposta em freqüência. Uma outra vantagem desta abordagem é que os testes de resposta em freqüência são, em geral, simples e podem ser realizados precisamente pelo uso de geradores de sinal senoidal realmente disponíveis e equipamentos de medida precisos. Comumente funções de transferência de componentes complicados são determinadas experimentalmente através de testes de resposta em freqüência. Além disso a abordagem da resposta em freqüência possui as vantagens de que um sistema pode ser projetado de modo que os efeitos de ruídos indesejáveis sejam desprezíveis e a análise e projeto podem ser estendidos a alguns sistemas de controle não lineares.

Embora a resposta em freqüência de um sistema de controle apresente uma imagem qualitativa da resposta transitória, a correlação entre as respostas em freqüência e transitória é indireta, exceto no caso de sistemas de segunda-ordem. No projeto de um sistema de malha-fechada, devemos ajustar a característica de resposta em freqüência usando alguns critérios de projeto (a serem apresentados na Seção 9.7) de modo a obter características de resposta transitória aceitáveis.

Uma vez entendida a correlação indireta entre algumas medidas da resposta transitória e da resposta em freqüência, a abordagem de resposta em freqüência pode ser utilizada com vantagem. O projeto de um sistema de controle através desta abordagem é baseado na interpretação das características dinâmicas desejadas em termos das características de resposta em freqüência. Esta análise de um sistema de controle indica graficamente quais variações ou modificações devem-se fazer na função de transferência de malha-aberta a fim de se obter as características de resposta transitória desejadas.

Obtenção de soluções em regime estacionário para entradas senoidais. Provaremos inicialmente o fato básico de que as características de resposta em freqüência de um sistema podem ser obtidas diretamente da função de transferência senoidal, isto é, a função de transferência na qual s é substituída por $j\omega$, onde ω é a freqüência angular.

Considere o sistema linear invariante no tempo indicado na Fig. 9.1. A entrada e a saída do sistema, cuja função de transferência é $G(s)$, são denotadas por $x(t)$ e $y(t)$, respectivamente. A entrada $x(t)$ é senoidal e dada por

$$x(t) = X \operatorname{sen} \omega t$$

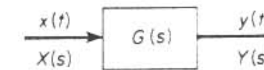


Fig. 9.1 Sistema linear invariante no tempo.

Suponha que a função de transferência $G(s)$ pode ser escrita como uma relação de dois polinômios em s; isto é,

$$G(s) = \frac{p(s)}{q(s)} = \frac{p(s)}{(s+s_1)(s+s_2)\cdots(s+s_n)}$$

A transformada de Laplace da saída $Y(s)$ é então

$$Y(s) = G(s)X(s) = \frac{p(s)}{q(s)}X(s) \tag{9-1}$$

onde $X(s)$ é a transformada de Laplace da entrada $x(t)$.

Limitaremos nossa discussão apenas a sistemas estáveis. Para estes sistemas, as partes reais das raízes $-s_i$ são negativas. A resposta de regime estacionário de um sistema linear, invariante no tempo e estável, para uma entrada senoidal não depende das condições iniciais. (Portanto, podemos supor condições iniciais nulas.)

Se $Y(s)$ possuir apenas pólos distintos, então a expansão em frações parciais da Eq. (9.1) fornece

$$\begin{aligned} Y(s) &= \frac{p(s)}{q(s)}\frac{\omega X}{s^2+\omega^2} \\ &= \frac{a}{s+j\omega} + \frac{\bar{a}}{s-j\omega} + \frac{b_1}{s+s_1} + \frac{b_2}{s+s_2} + \cdots + \frac{b_n}{s+s_n} \end{aligned} \tag{9-2}$$

onde a e b_i (onde $i = 1, 2, \ldots, n$) são constantes e $\bar{a}$ é o complexo conjugado de a. A transformada de Laplace inversa da Eq. (9.2) resulta em

$$y(t) = ae^{-j\omega t} + \bar{a}e^{j\omega t} + b_1e^{-s_1t} + b_2e^{-s_2t} + \cdots + b_ne^{-s_nt} \qquad (t \geq 0) \tag{9-3}$$

Para um sistema estável, $-s_1, -s_2, \ldots -s_n$ possuem partes reais negativas. Portanto, conforme t tende a infinito, os termos $e^{-s_1 t}, e^{-s_2 t}, \ldots$, e $s^{-s_n t}$ tendem a zero. Portanto, todos os termos do segundo membro da Eq. (9.3), exceto os dois primeiros, se anulam em regime estacionário.

Se $Y(s)$ envolve pólos múltiplos s_j de multiplicidade m_j, então $y(t)$ envolverá termos do tipo $t^{h_j}e^{-s_j t}(h_j = 0, 1, 2, \ldots, m_j - 1)$. Desde que as partes reais das raízes $-s_j$ são negativas para um sistema estável, os termos $t^{h_j}e^{-s_j t}$ tendem a zero conforme t tende a infinito.

Portanto, independentemente de o sistema possuir pólos distintos ou não, a resposta em regime estacionário resulta em

$$y(t) = ae^{-j\omega t} - \bar{a}e^{j\omega t} \tag{9-4}$$

onde a constante a pode ser calculada a partir da Eq. (9.2) como segue:

$$a = G(s)\frac{\omega X}{s^2 - \omega^2}(s + j\omega)\bigg|_{s=-j\omega} = -\frac{XG(-j\omega)}{2j}$$

Note que

$$\bar{a} = G(s)\frac{\omega X}{s^2 - \omega^2}(s - j\omega)\bigg|_{s=j\omega} = \frac{XG(j\omega)}{2j}$$

Desde que $G(j\omega)$ é uma função complexa, pode ser escrita na seguinte forma:

$$G(j\omega) = |G(j\omega)|\,e^{j\phi}$$

onde $|G(j\omega)|$ representa o módulo e ϕ representa o ângulo de $G(j\omega)$; isto é,

$$\phi = \underline{/G(j\omega)} = \text{tg}^{-1}\left[\frac{\text{parte imaginária de } G(j\omega)}{\text{parte real de } G(j\omega)}\right]$$

O ângulo ϕ pode ser negativo, positivo ou nulo. Analogamente, para $G(-j\omega)$ obtemos a seguinte expressão:

$$G(-j\omega) = |G(-j\omega)|\,e^{-j\phi} = |G(j\omega)|\,e^{-j\phi}$$

Então, a Eq. (9.4) pode ser escrita

$$\begin{aligned} y(t) &= X|G(j\omega)|\frac{e^{j(\omega t+\phi)} - e^{-j(\omega t+\phi)}}{2j} \\ &= X|G(j\omega)|\,\text{sen}\,(\omega t + \phi) \\ &= Y\,\text{sen}\,(\omega t + \phi) \end{aligned} \tag{9-5}$$

onde $Y = X\,|G(j\omega)|$. Verificamos que um sistema linear, estável, invariante no tempo e sujeito a uma entrada senoidal possuirá, em regime estacionário, uma saída senoidal com a mesma freqüência da entrada. Porém, a amplitude e o ângulo de fase da saída, em geral, serão diferentes daqueles da entrada. De fato, a amplitude da saída é dada pelo produto da amplitude da entrada por $|G(j\omega)|$, enquanto o ângulo de fase da saída difere daquele da entrada de um valor $\phi = \underline{/G(j\omega)}$. Um exemplo de entrada e saída com sinais senoidais é indicado na Fig. 9.2.

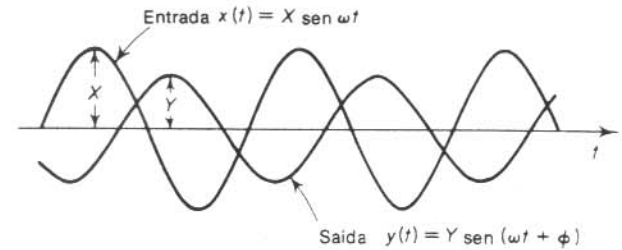


Fig. 9.2 Sinais senoidais de entrada e saída.

Com base no exposto, obtemos este importante resultado: Para entradas senoidais,

$$|G(j\omega)| = \left|\frac{Y(j\omega)}{X(j\omega)}\right| = \text{relação de amplitude entre a saída senoidal e a entrada senoidal}$$

$$\underline{/G(j\omega)} = \underline{\Big/\frac{Y(j\omega)}{X(j\omega)}} = \text{defasagem da saída senoidal em relação à entrada senoidal}$$

Portanto, as características de resposta de um sistema para entrada senoidal podem ser obtidas diretamente de

$$\frac{Y(j\omega)}{X(j\omega)} = G(j\omega)$$

A função de transferência senoidal $G(j\omega)$, a relação entre $Y(j\omega)$ e $X(j\omega)$, é uma função complexa e pode ser representada pelo módulo e ângulo de fase, tendo a freqüência como variável ou parâmetro. (Um ângulo de fase negativo é denominado *atraso de fase* e um ângulo de fase positivo é denominado *avanço de fase*.) A função de transferência senoidal de qualquer sistema linear é obtida substituindo-se s por $j\omega$ na função de transferência do sistema. Para caracterizar completamente um sistema linear no domínio de freqüência, devemos especificar tanto a relação de amplitude como o ângulo de fase como funções da freqüência ω.

Exemplo 9.1 Considere o sistema indicado na Fig. 9.3. A função de transferência $G(s)$ é

$$G(s) = \frac{K}{Ts + 1}$$

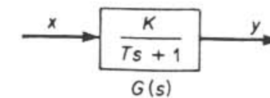


Fig. 9.3 Sistema de primeira-ordem.

Para a entrada senoidal $x(t) = X \operatorname{sen} \omega t$, a saída $y(t)$ pode ser determinada como segue. Substituindo-se s por $j\omega$ em $G(s)$, resulta

$$G(j\omega) = \frac{K}{jT\omega + 1}$$

A relação entre as amplitudes da saída e da entrada é

$$|G(j\omega)| = \frac{K}{\sqrt{1 + T^2\omega^2}}$$

enquanto o ângulo de fase ϕ é

$$\phi = \underline{/G(j\omega)} = -\operatorname{tg}^{-1} T\omega$$

Portanto, para a entrada $x(t) = X \operatorname{sen} \omega t$, a saída $y(t)$ pode ser obtida a partir da Eq. (9.5) como segue:

$$y(t) = \frac{XK}{\sqrt{1 + T^2\omega^2}} \operatorname{sen}(\omega t - \operatorname{tg}^{-1} T\omega) \tag{9-6}$$

Da Eq. (9.6) observa-se que para ω pequeno, a amplitude da saída $y(t)$ é praticamente igual a K vezes a amplitude da entrada. A defasagem da saída é pequena para valores pequenos de ω. Para ω grande, a amplitude da saída é pequena e praticamente inversamente proporcional a ω. A defasagem tende a $-90°$ conforme ω tende a infinito.

Resposta em freqüência a partir de diagramas de pólos e zeros. A resposta em freqüência pode ser determinada graficamente a partir do diagrama de pólos e zeros da função de transferência. Considere a seguinte função de transferência:

$$G(s) = \frac{K(s + z)}{s(s + p)}$$

onde p e z são reais. A resposta em freqüência desta função de transferência pode ser obtida de

$$G(j\omega) = \frac{K(j\omega + z)}{j\omega(j\omega + p)}$$

Os fatores $j\omega + z$, $j\omega$ e $j\omega + p$ são números complexos, conforme indicado na Fig. 9.4. O módulo de $G(j\omega)$ é

$$|G(j\omega)| = \frac{K|j\omega + z|}{|j\omega||j\omega + p|} = \frac{K|\overline{AP}|}{|\overline{OP}| \cdot |\overline{BP}|}$$

e o ângulo de fase de $G(j\omega)$ é

$$\begin{aligned} \underline{/G(j\omega)} &= \underline{/j\omega + z} - \underline{/j\omega} - \underline{/j\omega + p} \\ &= \operatorname{tg}^{-1} \frac{\omega}{z} - 90° - \operatorname{tg}^{-1} \frac{\omega}{p} \\ &= \phi - \theta_1 - \theta_2 \end{aligned}$$

onde os ângulos ϕ e θ_1 e θ_2 são definidos na Fig. 9.4. Note que uma rotação no sentido anti-horário é definida como o sentido positivo para a medida de ângulo.

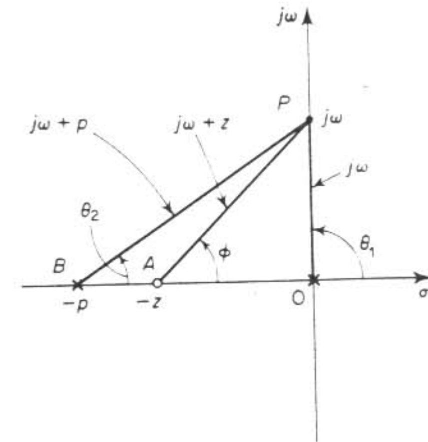


Fig. 9.4 Determinação da resposta em freqüência no plano complexo.

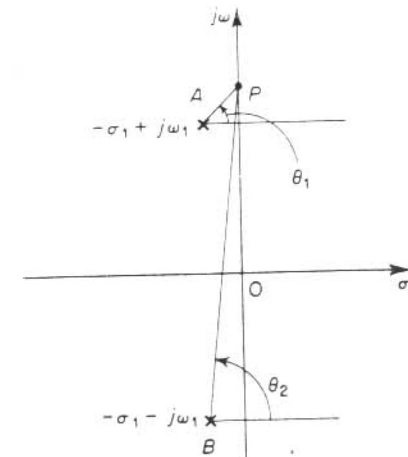


Fig. 9.5 Determinação da resposta em freqüência no plano complexo.

A partir da análise de resposta transitória de sistemas de malha-fechada, sabemos que um par de pólos complexos conjugados próximos ao eixo $j\omega$ produzirão um tipo altamente oscilatório de resposta transitória. No caso de resposta em freqüência, este par de pólos produzirá uma resposta com um pico altamente significativo.

Considere, por exemplo, a seguinte função de transferência:

$$G(s) = \frac{K}{(s + p_1)(s + p_2)}$$

onde p_1 e p_2 são complexos conjugados, conforme indicação na Fig. 9.5. A resposta em freqüência desta função de transferência pode ser determinada de

$$|G(j\omega)| = \frac{K}{|j\omega + p_1||j\omega + p_2|}$$

$$= \frac{K}{|\overline{AP}||\overline{BP}|}$$

$$\underline{/G(j\omega)} = -\theta_1 - \theta_2$$

onde os ângulos θ_1 e θ_2 são definidos na Fig. 9.5. Desde que $|\overline{AP}||\overline{BP}|$ é muito pequeno próximo a $\omega = \omega_1$, $|G(j\omega_1)|$ é muito grande. Portanto um par de pólos complexos conjugados próximos ao eixo $j\omega$ acarretará uma resposta em freqüência com grande valor de pico.

Inversamente, se a resposta em freqüência não apresentar um pico significativo, então a função de transferência não possui pólos complexos conjugados próximos ao eixo $j\omega$. Esta função de transferência também não exibirá uma resposta transitória altamente oscilatória. Desde que a resposta em freqüência indiretamente descreve a localização dos pólos e zeros da função de transferência, podemos estimar as características de resposta transitória de um sistema a partir das características de resposta em freqüência. Na Seção 9.7 apresentaremos uma discussão detalhada deste assunto.

9.2 GRÁFICOS LOGARÍTMICOS

Esta seção, bem como as duas seguintes são relacionadas principalmente com a apresentação das características de resposta em freqüência de sistemas de controle lineares.

A função de transferência senoidal, uma função complexa da freqüência ω, é caracterizada pelo seu módulo e ângulo de fase, com a freqüência como parâmetro. Há três representações comumente utilizadas de funções de transferência senoidais. Elas são

1. gráfico logarítmico ou diagrama de Bode
2. gráfico polar
3. gráfico do log-módulo *versus* fase

Esta seção apresenta gráficos logarítmicos de funções de transferência senoidais. Gráficos polares e gráficos log-módulo *versus* fase são apresentados nas Seções 9.3 e 9.4, respectivamente.

Gráficos logarítmicos ou diagramas de Bode. Uma função de transferência senoidal pode ser representada por dois gráficos separados, um fornecendo o módulo *versus* freqüência e outro o ângulo de fase *versus* freqüência. Um gráfico logarítmico ou diagrama de Bode consiste em dois gráficos: Um deles é um gráfico do logaritmo do módulo de uma função de transferência senoidal: o outro é um gráfico do ângulo de fase; ambos são construídos em função da freqüência, esta em uma escala logarítmica.

A representação padrão do módulo logarítmico de $G(j\omega)$ é $20 \log |G(j\omega)|$, onde a base do logaritmo é 10. A unidade usada nesta representação do módulo é o decibel, usualmente abreviado db. Na representação logarítmica, as curvas são desenhadas em papel monolog, usando-se a escala log para a freqüência e a escala linear tanto para o módulo (porém em db) como para o ângulo de fase (em graus). (A faixa de freqüência de interesse determina o número de décadas logarítmicas exigidas na abscissa.)

A principal vantagem de usar gráfico logarítmico é que a multiplicação dos módulos é convertida em uma adição. Além disso, é disponível um método simples para esboçar uma curva aproximada do log-módulo. Esta é baseada em aproximações assintóticas. Estas aproximações por retas assintóticas são suficientes somente se for necessária uma informação grosseira das características de resposta em freqüência. Se forem desejadas as curvas exatas, facilmente podem ser feitas as correções nas curvas assintóticas básicas. As curvas de ângulo de fase podem ser facilmente desenhadas se for disponível um modelo para a curva do ângulo de fase de $1 + j\omega$.

Note que a determinação experimental de uma função de transferência pode ser realizada de modo simples se os dados de resposta em freqüência estão apresentados na forma de um gráfico logarítmico.

A representação logarítmica é útil pelo fato de que mostra tanto as características de baixa freqüência como aquelas de alta freqüência, para a função de transferência considerada, em um único diagrama. A expansão da faixa de baixa freqüência utilizando uma escala logarítmica para a freqüência é muito vantajosa desde que as características de baixa freqüência são mais importantes em sistemas práticos. (Note que devido à escala de freqüência logarítmica, é impossível construir as curvas para a direita até a freqüência zero; entretanto, isto não cria qualquer problema sério.)

Fatores básicos de $G(j\omega)H(j\omega)$. Conforme citado anteriormente, a principal vantagem em usar o gráfico logarítmico está na relativa facilidade de desenhar as curvas de resposta em freqüência. Os fatores básicos que mais freqüentemente ocorrem em uma função de transferência arbitrária $G(j\omega)H(j\omega)$ são:

1. ganho K
2. fatores integral e derivativo $(j\omega)^{\mp 1}$
3. fatores de primeira-ordem $(1 + j\omega T)^{\mp 1}$
4. fatores quadráticos $[1 + 2\zeta\,(j\omega/\omega_n) + (j\omega/\omega_n)^2]^{\mp 1}$

Uma vez familiarizados com os gráficos logarítmicos destes fatores básicos, é possível utilizá-los na construção de gráficos logarítmicos compostos para qualquer forma geral de $G(j\omega)H(j\omega)$ esboçando-se as curvas para cada fator e adicionando as curvas individuais graficamente, porque a adição dos logaritmos dos ganhos correspondem à multiplicação dos mesmos.

O processo de obtenção do gráfico logarítmico pode ser ainda mais simplificado pelo uso de aproximações assintóticas para as curvas de cada fator. (Se necessário, podem ser feitas correções facilmente em um gráfico aproximado para obter-se um gráfico preciso.)

O ganho K. Um número maior do que a unidade possui um valor positivo em decibéis, enquanto um número menor do que a unidade possui um valor negativo. A curva do log-módulo para um ganho constante K é uma horizontal de valor $20 \log K$

db. O ângulo de fase do ganho K é nulo.* O efeito da variação do ganho K na função de transferência é que ele desloca para cima ou para baixo a curva do log-módulo da função de transferência por uma quantidade correspondente constante, não afetando porém o ângulo de fase.

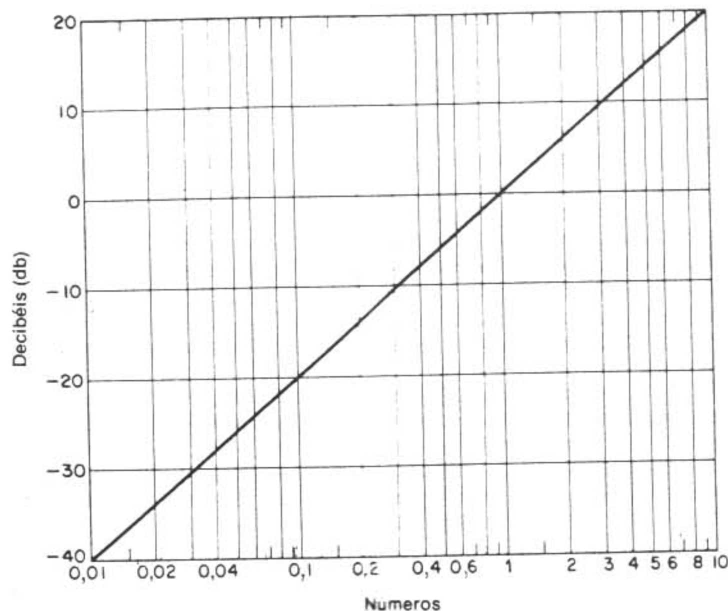


Fig. 9.6 Reta para conversão número-decibel.

Uma reta de conversão número-decibel é fornecida na Fig. 9.6. O valor em decibel de qualquer número pode ser obtido a partir desta reta. Quando um número aumenta por um fator de 10, o valor em decibel correspondente aumenta por um fator de 20. Este resultado pode ser verificado a partir do seguinte:

$$20 \log (K \times 10^n) = 20 \log K + 20n$$

Note que, quando expresso em db, o recíproco de um número difere do seu valor apenas no sinal; isto é, para um número K,

$$20 \log K = -20 \log \frac{1}{K}$$

Fatores integral e derivativo $(j\omega)^{\mp 1}$. O módulo logarítmico de $1/j\omega$ em db é

$$20 \log \left| \frac{1}{j\omega} \right| = -20 \log \omega \text{ db}$$

O ângulo de fase de $1/j\omega$ é uma constante e igual a -90°.

*N. do T. O autor está considerando $K > 0$.

Em gráficos logarítmicos, as relações de freqüências são expressas em termos de oitavas ou décadas. Uma oitava é um intervalo de freqüência desde ω_1 até $2\omega_1$, onde ω_1 é uma freqüência de qualquer valor. Uma década corresponde a um intervalo de freqüência desde ω_1 até $10\omega_1$, onde, novamente, ω_1 é qualquer freqüência. (Na escala logarítmica de um papel monolog qualquer relação de freqüência pode ser representada pela mesma distância horizontal. Por exemplo, a distância horizontal desde $\omega = 1$ até $\omega = 10$ é igual à distância desde $\omega = 3$ até $\omega = 30$.)

Se o log-módulo $-20 \log \omega$ db é colocado em um gráfico com ω na escala logarítmica, a curva resultante é uma reta. Desde que

$$(-20 \log 10\omega) \text{ db} = (-20 \log \omega - 20) \text{ db}$$

a inclinação da reta é -20 db/década (ou -6 db/oitava).

Analogamente, o log do módulo de $j\omega$ em db é

$$20 \log |j\omega| = 20 \log \omega \text{ db}$$

O ângulo de fase de $j\omega$ é constante e igual a 90°. A curva do log-módulo é uma reta com inclinação de 20 db/década. As Figs. 9.7(a) e (b) mostram as curvas de resposta em freqüência para $1/j\omega$ e $j\omega$, respectivamente. Podemos verificar facilmente que as diferenças nas curvas de resposta em freqüência dos fatores $1/j\omega$ e $j\omega$ correspondem aos sinais das inclinações das curvas do log-módulo e aos sinais dos ângulos de fase. Ambas as curvas dos logs dos módulos tornam-se iguais a 0 db para $\omega = 1$.

Se a função de transferência contém o fator $(1/j\omega)^n$ ou $(j\omega)^n$, o log do módulo

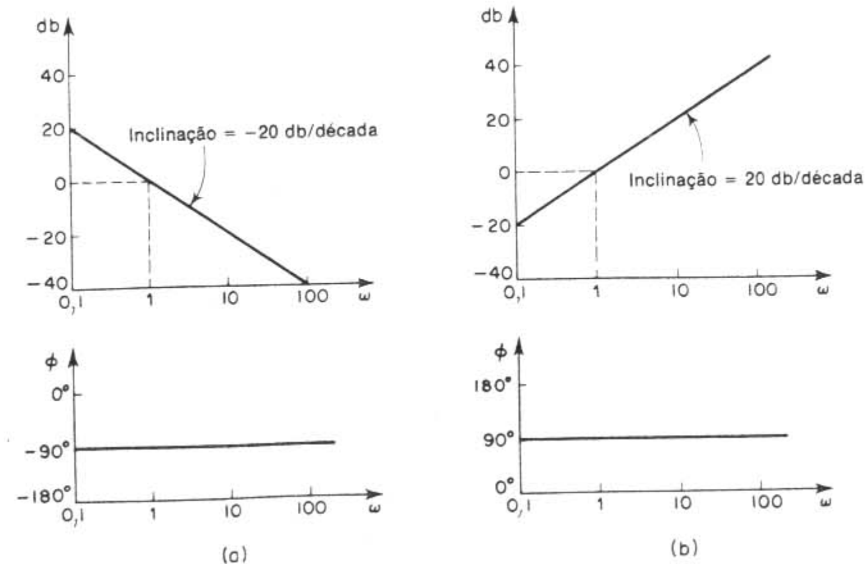


Fig. 9.7 (a) Curvas de resposta em freqüência de $1/j\omega$; (b) curvas de resposta em freqüência de $j\omega$.

resulta, respectivamente, em

$$20 \log \left| \frac{1}{(j\omega)^n} \right| = -n \times 20 \log |j\omega| = -20n \log \omega \text{ db}$$

ou

$$20 \log |(j\omega)^n| = n \times 20 \log |j\omega| = 20n \log \omega \text{ db}$$

As inclinações das curvas dos log-módulos para os fatores $(1/j\omega)^n$ e $(j\omega)^n$ são então $-20n$ db/década e $20n$ db/década, respectivamente. O ângulo de fase de $(1/j\omega)^n$ é igual a $-90° \times n$ em toda a faixa de freqüência, enquanto que para $(j\omega)^n$ é igual a $90° \times n$ em toda a faixa de freqüência.

Fatores de primeira-ordem $(1 + j\omega T)^{\mp 1}$. O log do módulo do fator de primeira-ordem $1/(1 + j\omega T)$ é

$$20 \log \left| \frac{1}{1 + j\omega T} \right| = -20 \log \sqrt{1 + \omega^2 T^2} \text{ db}$$

Para baixas freqüências, tais como $\omega << 1/T$, o log do módulo pode ser aproximado por

$$-20 \log \sqrt{1 + \omega^2 T^2} \doteqdot -20 \log 1 = 0 \text{ db}$$

Portanto, a curva do log-módulo em baixas freqüências é a reta constante 0-db. Para altas freqüências, tais como $\omega >> 1/T$,

$$-20 \log \sqrt{1 + \omega^2 T^2} \doteqdot -20 \log \omega T \text{ db}$$

Esta é uma expressão aproximada para a faixa de altas freqüências. Em $\omega = 1/T$, o log do módulo é igual a 0 db; em $\omega = 10/T$, o log do módulo é -20 db. Portanto, o valor de $-20 \log \omega T$ db decresce de 20 db para cada década de ω. Para $\omega >> 1/T$, a curva do log do módulo é uma reta com inclinação de -20 db/década (ou -6 db/oitava).

A análise anterior mostra que a representação logarítmica da curva de resposta em freqüência do fator $1/(1 + j\omega T)$ pode ser aproximada por duas retas assintóticas, uma reta em 0 db para a faixa de freqüência $0 < \omega < 1/T$, e uma outra reta com inclinação -20 db/década (ou -6 db/oitava) para a faixa de freqüência $1/T < \omega < \infty$. A curva exata do log-módulo, as assíntotas, e a curva exata do ângulo de fase são mostradas na Fig. 9.8.

A freqüência na qual as duas assíntotas se interceptam é denominada freqüência de *canto* ou freqüência de *mudança de inclinação* ("*quebra*"). Para o fator $1/(1 + j\omega T)$, a freqüência $\omega = 1/T$ é a freqüência de canto desde que em $\omega = 1/T$ as duas assíntotas possuem o mesmo valor. (A expressão assintótica em baixa-freqüência em $\omega = 1/T$ é 20 log 1 db = 0 db e a expressão assintótica de alta-freqüência em $\omega = 1/T$ é também 20 log 1 db = 0 db.) A freqüência de canto divide a curva de resposta em freqüência em duas regiões, uma curva para a região de baixa-freqüência e uma curva para a região de alta-freqüência. A freqüência de canto é muito importante para o esboço de curvas de resposta em freqüência logarítmicas.

O ângulo de fase exato ϕ do fator $1/(1 + j\omega T)$ é

$$\phi = -\text{tg}^{-1} \ \omega T$$

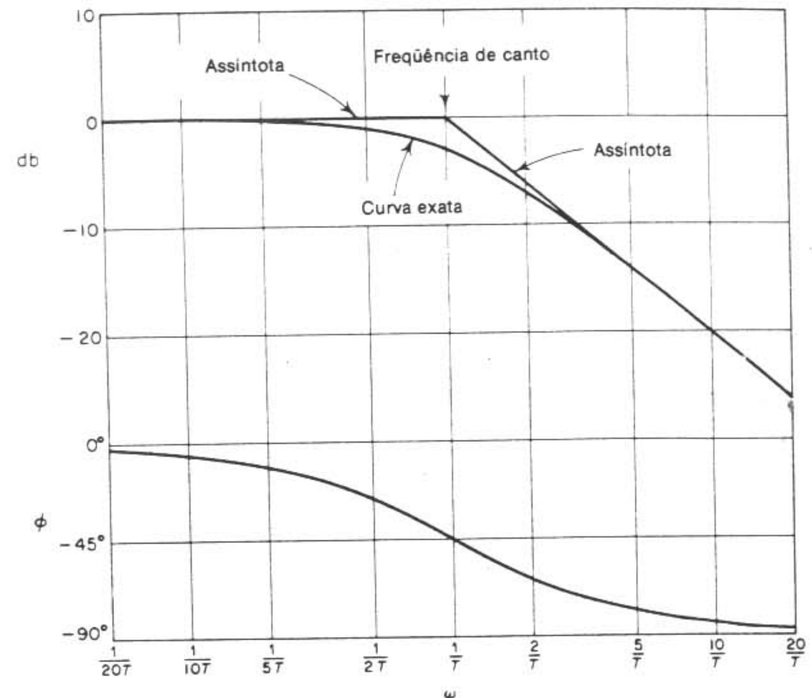


Fig. 9.8 Curva do log-módulo conjuntamente com as assíntotas e curvas do ângulo de fase de $1/(1 + j\omega T)$.

Na freqüência zero, o ângulo de fase é 0°. Na freqüência de canto, o ângulo de fase é

$$\phi = -\text{tg}^{-1} \ \frac{T}{T} = -\text{tg}^{-1} \ 1 = -45°$$

No infinito, o ângulo de fase torna-se $-90°$. Desde que o ângulo de fase é dado por uma função inversa da tangente, o ângulo de fase é anti-simétrico em relação ao ponto de inflexão em $\phi = -45°$.

O erro na curva de módulo causado pelo uso de assíntotas pode ser calculado. O erro máximo ocorre na freqüência de canto e é aproximadamente igual a -3 db desde que

$$-20 \log \sqrt{1 + 1} + 20 \log 1 = -10 \log 2 = -3{,}03 \text{ db}$$

O erro na freqüência uma oitava abaixo da freqüência de canto, isto é, em $\omega = 1/2T$, é

$$-20 \log \sqrt{\frac{1}{4} + 1} + 20 \log 1 = -20 \log \frac{\sqrt{5}}{2} = -0{,}97 \text{ db}$$

O erro na freqüência uma oitava acima da freqüência de canto, isto é, em $\omega = 2/T$, é

$$-20 \log \sqrt{2^2 + 1} + 20 \log 2 = -20 \log \frac{\sqrt{5}}{2} = -0{,}97 \text{ db}$$

Portanto, o erro em uma oitava abaixo ou acima da freqüência de canto é aproximadamente igual a -1 db. Analogamente, o erro em uma década abaixo ou acima da freqüência de canto é aproximadamente $-0{,}04$ db. O erro em decibéis envolvido no uso da expressão assintótica para a curva de resposta em freqüência de $1/(1 + j\omega T)$ é indicado na Fig. 9.9. O erro é simétrico em relação à freqüência de canto.

Desde que as assíntotas são muito fáceis de desenhar e suficientemente próximas da curva exata, o uso destas aproximações nos desenhos dos diagramas de Bode é conveniente para estabelecer a natureza geral das características de resposta em freqüência rapidamente e com um mínimo de cálculo, podendo ser utilizado para muitos trabalhos de projeto preliminares. Se forem desejadas as curvas precisas de resposta em freqüência, podem ser feitas as correções facilmente, referindo-se à curva fornecida na Fig. 9.9. Na prática, uma curva precisa de resposta em freqüência pode ser desenhada localizando-se o ponto de -3 db na freqüência de canto e os pontos de -1 db uma oitava acima e abaixo da freqüência de canto e então ligando estes pontos por uma curva suave.

Note que uma variação na constante de tempo T desloca a freqüência de canto para a direita ou para a esquerda, porém as formas das curvas do log-módulo e do ângulo de fase permanecem as mesmas.

A função de transferência $1/(1 + j\omega T)$ possui as características de um filtro passa-baixas. Para freqüências acima de $\omega = 1/T$, o log do módulo cai rapidamente para $-\infty$. Este resultado é devido essencialmente à presença da constante de tempo. No filtro passa-baixas, a saída pode seguir uma entrada senoidal muito precisamente em baixas freqüências. Porém, conforme a freqüência de entrada é aumentada, a saída não pode seguir a entrada porque é necessário um certo intervalo de tempo para o sistema atingir o valor necessário. Portanto, em altas freqüências, a amplitude da saída tende a zero e o ângulo de fase da saída tende a $-90°$. Portanto, se a função de entrada contém muitas harmônicas, então as componentes de baixas freqüências são bem reproduzidas na saída, enquanto as componentes de altas freqüências são atenuadas em amplitude e defasadas. Portanto, um

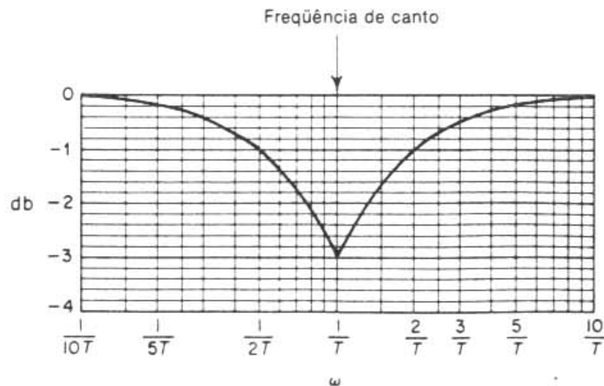


Fig. 9.9 Erro do log-módulo na expressão assintótica da curva de resposta em freqüência de $1/(1 + j\omega T)$.

elemento de primeira-ordem fornece uma réplica exata, ou quase exata, apenas para fenômenos constantes ou lentamente variáveis.

Uma vantagem da representação logarítmica é que para fatores recíprocos, por exemplo, o fator $1 + j\omega T$, as curvas do log-módulo e do ângulo de fase necessitam apenas trocar de sinal. Desde que

$$20 \log |1 + j\omega T| = -20 \log \left| \frac{1}{1 + j\omega T} \right|$$

$$\underline{/1 + j\omega T} = \text{tg}^{-1}\ \omega T = -\underline{\Bigg/\frac{1}{1 + j\omega T}}$$

a freqüência de canto é a mesma em ambos os casos. A inclinação da assíntota de alta freqüência de $1 + j\omega T$ é 20 db/década, e o ângulo de fase varia desde 0° até 90° conforme a freqüência ω aumenta desde 0 até infinito. A curva do log-módulo, juntamente com as assíntotas e a curva ângulo de fase para o fator $1 + j\omega T$ são indicadas na Fig. 9.10.

Os formatos das curvas do ângulo de fase são os mesmos para qualquer fator da forma $(1 + j\omega T)^{\mp 1}$. Portanto, é conveniente possuir um modelo para a curva do ângulo de fase à disposição. Este modelo deve ser usado repetidamente para construção das curvas do ângulo de fase para qualquer função da forma $(1 + j\omega T)^{\mp 1}$. Se este modelo não for disponível, devemos localizar alguns pontos da curva. Os ângulos de fase de $(1 + j\omega T)^{\mp 1}$ são

$$\mp 45° \quad \text{em} \quad \omega = \frac{1}{T}$$

$$\mp 26{,}6° \quad \text{em} \quad \omega = \frac{1}{2T}$$

$$\mp 5{,}7° \quad \text{em} \quad \omega = \frac{1}{10T}$$

$$\mp 63{,}4° \quad \text{em} \quad \omega = \frac{2}{T}$$

$$\mp 84{,}3° \quad \text{em} \quad \omega = \frac{10}{T}$$

Para o caso no qual uma dada função de transferência envolve termos do tipo $(1 + j\omega T)^{\mp n}$, pode ser feita uma construção assintótica similar. A freqüência de canto ainda é em $\omega = 1/T$, e as assíntotas são retas. A assíntota de baixa freqüência é uma reta horizontal em 0 db, enquanto a assíntota de alta freqüência possui a inclinação $-20n$ db/década ou $20n$ db/década. O erro envolvido nas expressões assintóticas é n vezes o correspondente a $(1 + j\omega T)^{\mp 1}$. O ângulo de fase é n vezes aquele de $(1 + j\omega T)^{\mp 1}$ em cada freqüência.

Fatores quadráticos $[1 + 2\zeta(j\omega/\omega_n) + (j\omega/\omega_n)^2]^{\mp 1}$. Sistemas de controle normalmente possuem fatores quadráticos da forma

$$\frac{1}{1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2} \tag{9-7}$$

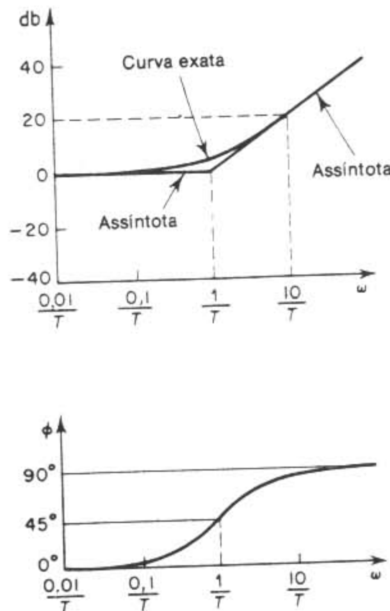


Fig. 9.10 Curva do log-módulo conjuntamente com as assíntotas e curvas do ângulo de fase de $1 + j\omega T$.

Se $\zeta > 1$, este fator quadrático pode ser expresso como um produto de dois fatores de primeira-ordem com pólos reais. Se $0 < \zeta < 1$, este fator quadrático é o produto de dois fatores complexos conjugados. As aproximações assintóticas para as curvas de resposta em freqüência não são precisas para um fator com baixos valores de ζ, pois tanto o módulo como o ângulo de fase do fator quadrático dependem da freqüência de canto e da relação de amortecimento ζ.

A curva de resposta em freqüência assintótica pode ser obtida como segue: Desde que

$$20 \log \left| \frac{1}{1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2} \right|$$

$$= -20 \log \sqrt{\left(1 - \frac{\omega^2}{\omega_n^2}\right)^2 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2}$$

para baixas freqüências tais que $\omega << \omega_n$, o log do módulo resulta em

$$-20 \log 1 = 0 \text{ db}$$

A assíntota de baixa freqüência é portanto uma reta horizontal em 0 db. Para altas freqüências tais que $\omega \gg \omega_n$, o log do módulo resulta em

$$-20 \log \frac{\omega^2}{\omega_n^2} = -40 \log \frac{\omega}{\omega_n} \text{ db}$$

A equação para a assíntota de alta freqüência é uma reta possuindo a inclinação −40 db/década desde que

$$-40 \log \frac{10\omega}{\omega_n} = -40 - 40 \log \frac{\omega}{\omega_n}$$

A assíntota de alta freqüência intercepta a de baixa freqüência em $\omega = \omega_n$ desde que nesta freqüência

$$-40 \log \frac{\omega}{\omega_n} = -40 \log 1 = 0 \text{ db}$$

Esta freqüência é a freqüência de canto do fator quadrático considerado.

As duas assíntotas que acabamos de deduzir são independentes do valor de ζ. Próximo à freqüência $\omega = \omega_n$, ocorre um pico de ressonância, conforme pode ser esperado a partir de (9.7). A relação de amortecimento ζ determina a amplitude deste pico de ressonância. Portanto, existem obviamente erros na aproximação pelas assíntotas. O valor do erro depende do valor de ζ. Será grande para pequenos valores de ζ. A Fig. 9.11 fornece as curvas exatas do log do módulo conjuntamente com as assíntotas e as curvas do ângulo de fase para o fator quadrático dado em (9.7) para alguns valores de ζ. Se forem desejadas correções nas curvas assintóticas, a correção necessária em um número suficiente de freqüências pode ser obtida a partir da Fig. 9.11.

O ângulo de fase do fator quadrático $[1 + 2\zeta(j\omega/\omega_n) + (j\omega/\omega_n)^2]^{-1}$ é

$$\phi = \angle \frac{1}{1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2} = -\operatorname{tg}^{-1}\left[\frac{2\zeta\frac{\omega}{\omega_n}}{1 - \left(\frac{\omega}{\omega_n}\right)^2}\right] \tag{9-8}$$

O ângulo de fase é uma função tanto de ω como de ζ. Em $\omega = 0$, o ângulo de fase é igual a 0°. Na freqüência de canto $\omega = \omega_n$, o ângulo de fase é −90° independentemente de ζ, desde que

$$\phi = -\operatorname{tg}^{-1}\left(\frac{2\zeta}{0}\right) = -\operatorname{tg}^{-1} \infty = -90°$$

Em $\omega = \infty$, o ângulo de fase resulta −180°. A curva do ângulo de fase é anti-simétrica em relação ao ponto de inflexão, o ponto onde $\phi = -90°$.

As curvas de resposta em freqüência para o fator

$$1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2$$

podem ser obtidas simplesmente invertendo-se o sinal daquelas do log do módulo e das curvas do ângulo de fase do fator

$$\frac{1}{1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2}$$

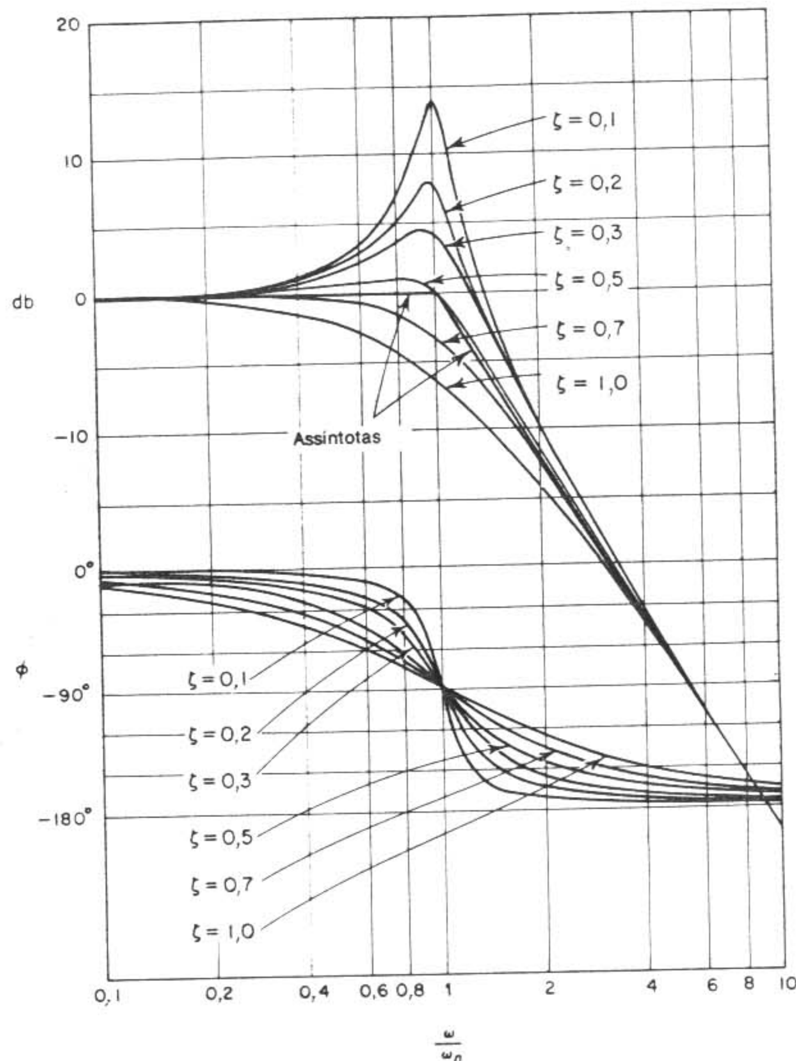


Fig. 9.11 Curvas do log-módulo conjuntamente com as assíntotas e curvas do ângulo de fase da função de transferência quadrática por (9.7).

Para obter as curvas de resposta em freqüência de uma dada função de transferência quadrática, devemos inicialmente determinar o valor da freqüência de canto ω_n e da relação de amortecimento ζ. Então, usando a família de curvas fornecida na Fig. 9.11, podem-se construir as curvas de resposta em freqüência.

A freqüência de ressonância ω_r e o valor do pico de ressonância M_r. O módulo de

$$G(j\omega) = \frac{1}{1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2}$$

é

$$|G(j\omega)| = \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_n^2}\right)^2 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2}} \tag{9-9}$$

Se $|G(j\omega)|$ possui um valor de pico em alguma freqüência, esta freqüência é denominada a freqüência de *ressonância*. Desde que o numerador de $|G(j\omega)|$ é constante, ocorrerá um valor de pico de $|G(j\omega)|$ quando

$$g(\omega) = \left(1 - \frac{\omega^2}{\omega_n^2}\right)^2 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2 \tag{9-10}$$

é um mínimo. Desde que a Eq. (9.10) pode ser escrita

$$g(\omega) = \left[\frac{\omega^2 - \omega_n^2(1 - 2\zeta^2)}{\omega_n^2}\right]^2 + 4\zeta^2(1 - \zeta^2) \tag{9-11}$$

o valor mínimo de $g(\omega)$ ocorre em $\omega = \omega_n\sqrt{1 - 2\zeta^2}$. Portanto, a freqüência de ressonância ω_r é

$$\omega_r = \omega_n\sqrt{1 - 2\zeta^2} \qquad (0 \le \zeta \le 0{,}707) \tag{9-12}$$

Conforme a relação de amortecimento ζ tende a zero, a freqüência de ressonância tende a ω_n. Para $0 < \zeta \le 0{,}707$, a freqüência de ressonância ω_r é menor do que a freqüência natural amortecida $\omega_d = \omega_n\sqrt{1 - \zeta^2}$, que é exibida na resposta transitória. Da Eq. (9.11) pode-se observar que, para $\zeta > 0{,}707$, não há pico de ressonância. Para $\zeta > 0{,}707$, o módulo $|G(j\omega)|$ decresce monotonicamente com o aumento na freqüência ω. Isto significa que não há pico na curva de módulo para $\zeta > 0{,}707$. (O módulo é menor do que 0 db para todos os valores de $\omega > 0$. Lembre-se que, para $0{,}7 < \zeta < 1$, a resposta ao degrau é oscilatória, porém as oscilações são bastante amortecidas e pouco perceptíveis.)

O valor do pico de ressonância M_r pode ser determinado substituindo-se a Eq. (9.12) na Eq. (9.9). Para $0 \le \zeta \le 0{,}707$,

$$M_r = |G(j\omega)|_{\text{máx.}} = |G(j\omega_r)| = \frac{1}{2\zeta\sqrt{1 - \zeta^2}} \tag{9-13}$$

Para $\zeta > 0{,}707$,

$$M_r = 1 \tag{9-14}$$

Conforme ζ tende a zero, M_r tende a infinito. Isto significa que se o sistema não amortecido for excitado em sua freqüência natural, o módulo de $G(j\omega)$ torna-se infinito. A relação entre M_r e ζ é indicada na Fig. 9.12.

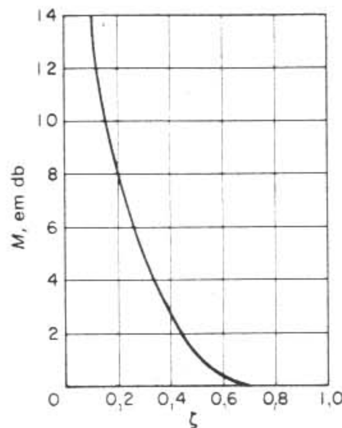


Fig. 9.12 Curva M_r *versus* ζ para o sistema de segunda-ordem $1/[1 + 2\zeta\,(j\omega/\omega_n) + (j\omega/\omega_n)^2]$.

O ângulo de fase de $G(j\omega)$ na freqüência onde ocorre o pico de ressonância pode ser obtido substituindo-se a Eq. (9.12) na Eq. (9.8). Portanto, na freqüência de ressonância ω_r,

$$\underline{/G(j\omega_r)} = -\,\text{tg}^{-1}\,\frac{\sqrt{1-2\zeta^2}}{\zeta}$$

$$= -90° + \text{sen}^{-1}\frac{\zeta}{\sqrt{1-\zeta^2}}$$

Procedimento geral para construção das curvas logarítmicas de resposta em freqüência. Inicialmente reescreva a função de transferência senoidal $G(j\omega)\,H(j\omega)$ como um produto dos fatores básicos anteriormente discutidos. Identifique então as freqüências de canto associadas com estes fatores básicos. Finalmente desenhe as curvas assintóticas do log-módulo, com inclinações apropriadas entre as freqüências de canto. A curva exata, situada muito próxima à curva assintótica, pode ser obtida efetuando-se as correções apropriadas.

A curva do ângulo de fase de $G(j\omega)H(j\omega)$ pode ser desenhada adicionando-se as curvas dos ângulos de fase dos fatores individuais.

O uso dos gráficos logarítmicos empregando aproximações assintóticas necessita um tempo muito menor do que os outros métodos que podem ser utilizados para determinação da resposta em freqüência de uma função de transferência. A facilidade de construção das curvas de resposta em freqüência para uma dada função de transferência e a facilidade de modificação da curva de resposta em freqüência, conforme é adicionada uma compensação, constituem as principais razões do uso muito comum na prática destes gráficos logarítmicos.

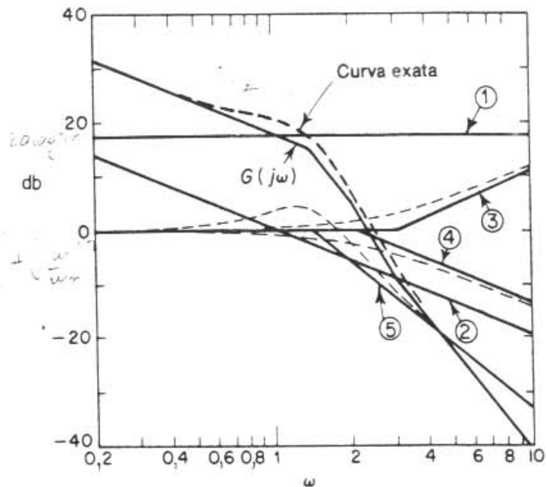
Exemplo 9.2 Desenhe um diagrama de blocos para a seguinte função de transferência:

$$G(j\omega) = \frac{10(j\omega + 3)}{(j\omega)(j\omega + 2)[(j\omega)^2 + j\omega + 2]}$$

Efetue correções de modo que a curva do log-módulo seja precisa.

De modo a evitar quaisquer erros possíveis na construção da curva do log-módulo, é desejável colocar $G(j\omega)$ na forma normalizada seguinte, onde as assíntotas de baixa freqüência para os fatores de primeira-ordem e para o fator de segunda-ordem correspondem à reta de 0-db.

$$G(j\omega) = \frac{7{,}5\left(\dfrac{j\omega}{3} + 1\right)}{(j\omega)\left(\dfrac{j\omega}{2} + 1\right)\left[\dfrac{(j\omega)^2}{2} + \dfrac{j\omega}{2} + 1\right]}$$



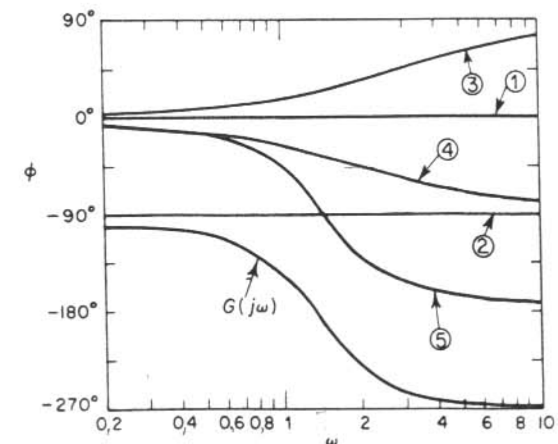


Fig. 9.13 Diagrama de Bode do sistema considerado no Exemplo 9.2.

Esta função é composta pelos seguintes fatores:

$$7{,}5, \quad (j\omega)^{-1}, \quad 1 + j\frac{\omega}{3}, \quad \left(1 + j\frac{\omega}{2}\right)^{-1} \quad \left[1 + j\frac{\omega}{2} + \frac{(j\omega)^2}{2}\right]^{-1}$$

As freqüências de canto do terceiro, quarto e quinto termos são $\omega = 3$, $\omega = 2$ e $\omega = \sqrt{2}$, respectivamente.

Para construir o diagrama de Bode, as curvas assintóticas separadas para cada um dos fatores são mostradas na Fig. 9.13. A curva composta é então obtida adicionando-se algebricamente as curvas individuais, também mostrada na Fig. 9.13. Note que quando as curvas assintóticas individuais são adicionadas em cada freqüência, a inclinação da curva composta é acumulativa. Abaixo de $\omega = \sqrt{2}$, o gráfico possui a inclinação de -20 db/década. Na primeira freqüência de canto $\omega = \sqrt{2}$, a inclinação muda para -60 db/década e continua com esta inclinação até a próxima freqüência de canto em $\omega = 2$, onde a inclinação torna-se -80 db/década. Na última freqüência de canto $\omega = 3$, a inclinação muda para -60 db/década.

Uma vez desenhada a curva aproximada do log-módulo, pode ser obtida a curva real adicionando-se as correções em cada freqüência de canto e nas freqüências uma oitava abaixo e acima das freqüências de canto. Para os fatores de primeira-ordem $(1 + j\omega T)^{\mp 1}$, as correções são ± 3 db na freqüência de canto e ± 1 db nas freqüências correspondentes a uma oitava acima e abaixo da freqüência de canto. As correções necessárias para o fator quadrático são obtidas da Fig. 9.11. A curva exata do log-módulo para $G(j\omega)$ é indicada por uma linha tracejada na Fig. 9.13.

Para a construção da curva completa do ângulo de fase, devem ser esboçadas as curvas do ângulo de fase para todos os fatores. A soma algébrica das curvas do ângulo de fase resulta na curva completa do ângulo de fase, conforme indicado na Fig. 9.13.

A simplicidade da construção do diagrama de Bode deve estar agora aparente. Após adquirir prática na construção de diagramas de Bode não é necessário construir as assíntotas para cada um dos fatores envolvidos na função de transferência. Podemos desenhar diretamente as curvas compostas, assintóticas, do log-módulo somando mentalmente as assíntotas para cada um dos fatores.

Sistemas de fase mínima e sistemas de fase não-mínima. Funções de transferência que não possuem pólos ou zeros no semiplano direito do plano s são funções de transferência de mínima fase; por outro lado, as funções de transferência que possuem pólos e/ou zeros no semiplano direito são funções de transferência de fase não-mínima. Os sistemas com funções de transferência com mínima fase são denominados sistemas de mínima fase; por outro lado, aqueles com funções de transferência de fase não-mínima são denominados sistemas de fase não-mínima.

Para sistemas com as mesmas características de módulo, a faixa no ângulo de fase da função de transferência de fase mínima é mínima para todos estes sistemas, enquanto a faixa no ângulo de fase de qualquer função de transferência de fase não-mínima é maior do que aquele mínimo.

Considere, como um exemplo, os dois sistemas cujas funções de transferência senoidais são, respectivamente,

$$G_1(j\omega) = \frac{1 + j\omega T}{1 + j\omega T_1}, \qquad G_2(j\omega) = \frac{1 - j\omega T}{1 + j\omega T_1} \qquad (0 < T < T_1)$$

As configurações de pólos e zeros destes sistemas são indicadas na Fig. 9.14. As duas funções de transferência senoidais possuem as mesmas características de módulo, porém as características do ângulo de fase são diferentes, conforme indicado na Fig. 9.15. Estes dois sistemas diferem entre si pelo fator

$$G(j\omega) = \frac{1 - j\omega T}{1 + j\omega T}$$

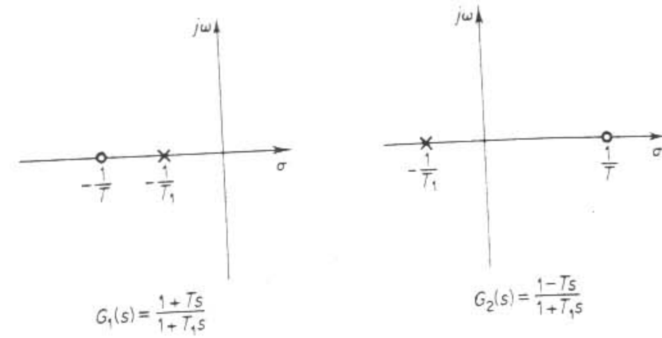


Fig. 9.14 Configurações de pólo-zero de um sistema de fase mínima $G_1(s)$ e um sistema de fase não-mínima $G_2(s)$.

O módulo do fator $(1 - j\omega T)/(1 + j\omega T)$ é sempre unitário. Porém, o ângulo de fase é igual a $-2\ \text{tg}^{-1}\,\omega T$ e varia desde 0 até $-180°$ conforme a freqüência varia desde zero até infinito.

Para um sistema de fase mínima, as características de módulo de ângulo e de fase são diretamente relacionadas. Isto significa que se a curva de módulo de um sistema for especificada em todas as freqüências desde zero até infinito, então a curva do ângulo de fase é unicamente determinada, e vice-versa. Isto, entretanto, não é verdade para um sistema de fase não-mínima.

As situações de fase não-mínima podem surgir através de dois modos diferentes. Um deles é simplesmente quando um sistema inclui um elemento, ou elementos, de fase não-mínima. A outra situação resulta nos casos onde um laço menor é instável.

Para um sistema de fase mínima, o ângulo de fase em $\omega = \infty$ é igual a $-90°(q - p)$, onde p e q são os graus dos polinômios do numerador e denominador da função de transferência, respectivamente. Para um sistema de fase não-mínima, o ângulo de fase em $\omega = \infty$ difere de $-90°(q - p)$. Em qualquer sistema, a inclinação da curva do log-módulo em $\omega = \infty$ é igual a $-20(q - p)$ db/década. É portanto possível detectar se o sistema é ou não de mínima fase examinando-se a inclinação da assíntota em alta freqüência da curva log-módulo e o ângulo de fase em $\omega = \infty$. Se a

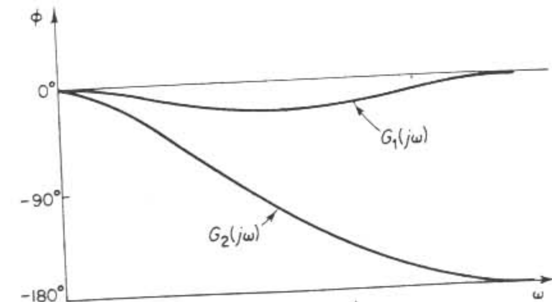


Fig. 9.15 Características do ângulo de fase dos sistemas $G_1(s)$ e $G_2(s)$ indicados na Fig. 9.14.

inclinação da curva do log-módulo, conforme a freqüência tende a infinito, for $-20(q - p)$ db/década e o ângulo de fase em $\omega = \infty$ é igual a $-90°(q - p)$, então o sistema é de mínima fase.

Sistemas de fase não-mínima são lentos em resposta em virtude de seus comportamentos defeituosos no início da resposta. Em muitos sistemas de controles práticos, devem ser cuidadosamente evitados atrasos de fase excessivos. No projeto de um sistema, se a velocidade rápida de resposta é de fundamental importância, não deveremos usar componentes de fase não-mínima. (Um exemplo comum de elementos de fase não-mínima que podem estar presentes em sistemas de controle é o atraso de transporte.)

Deve-se notar que as técnicas de análise e projeto de resposta em freqüência a serem apresentadas neste e no próximo capítulo são válidas tanto para sistemas de fase mínima como para sistemas de fase não-mínima.

Atraso de transporte. O atraso de transporte é um comportamento de fase não-mínima e possui um atraso de fase excessivo sem atenuação em altas freqüências. Considere o atraso de transporte dado por

$$G(j\omega) = e^{-j\omega T}$$

O módulo é sempre igual à unidade desde que

$$|G(j\omega)| = |\cos \omega T - j \operatorname{sen} \omega T| = 1$$

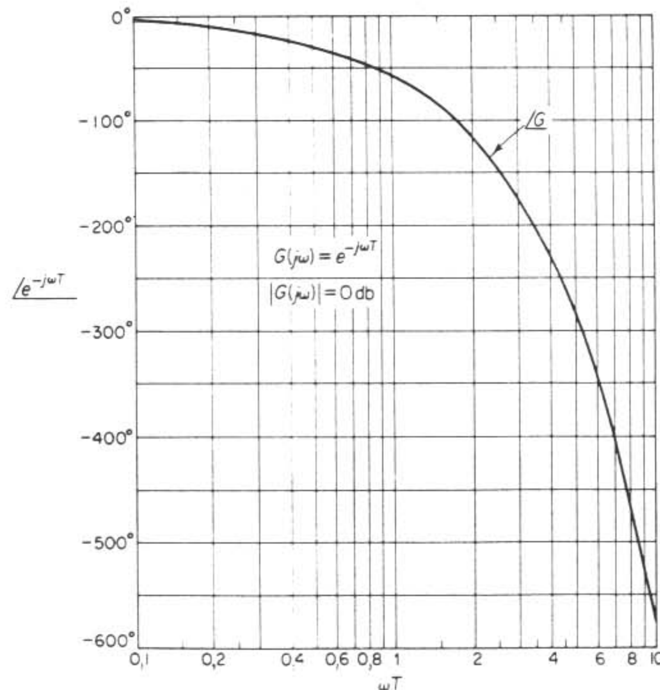


Fig. 9.16 Característica do ângulo de fase do atraso de transporte.

Portanto, o log do módulo do atraso de transporte $e^{-j\omega T}$ é igual a 0 db. O ângulo de fase do atraso de transporte é

$$\begin{aligned} \angle G(j\omega) &= -\omega T \qquad \text{(radianos)} \\ &= -57{,}3\ \omega T \qquad \text{(graus)} \end{aligned}$$

O ângulo de fase varia linearmente com a freqüência ω. A característica do ângulo de fase do atraso de transporte é indicada na Fig. 9.16.

Exemplo 9.3 Desenhe o diagrama de Bode da seguinte função de transferência:

$$G(j\omega) = \frac{e^{-j\omega L}}{1 + j\omega T}$$

O log do módulo é

$$\begin{aligned} 20 \log |G(j\omega)| &= 20 \log |e^{-j\omega L}| + 20 \log \left| \frac{1}{1 + j\omega T} \right| \\ &= 0 + 20 \log \left| \frac{1}{1 + j\omega T} \right| \end{aligned}$$

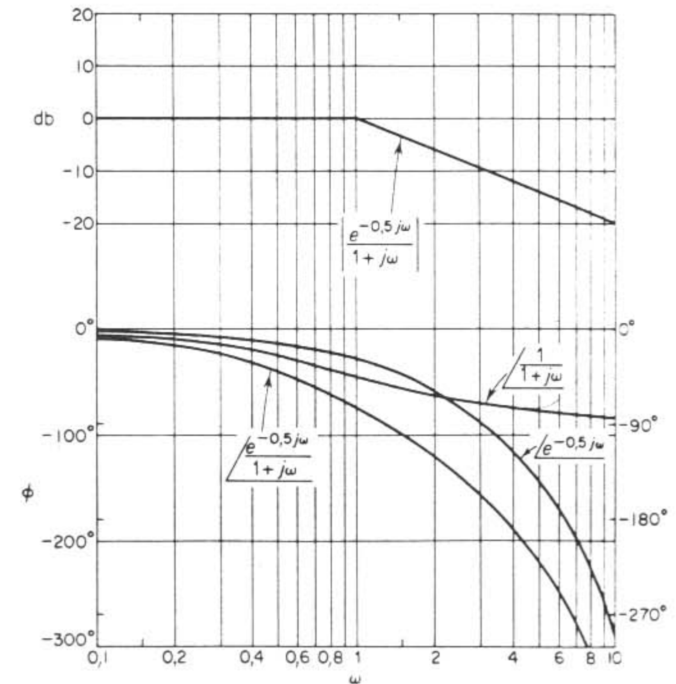


Fig. 9.17 Diagrama de Bode para o sistema $e^{-j\omega L}/(1 + j\omega T)$ com $L = 0{,}5$ e $T = 1$.

O ângulo de fase de $G(j\omega)$ é

$$\begin{aligned}\underline{/G(j\omega)} &= \underline{/e^{-j\omega L}} - \underline{/\frac{1}{1+j\omega T}} \\ &= -\omega L - \text{tg}^{-1}\ \omega T\end{aligned}$$

As curvas do log-módulo e ângulo de fase para esta função de transferência com $L = 0.5$ e $T = 1$ são mostradas na Fig. 9.17.

Relação entre o tipo do sistema e a curva do log-módulo. Os coeficientes de erro de posição, velocidade e aceleração estáticos descrevem o comportamento em baixa-freqüência dos sistemas tipo 0, tipo 1 e tipo 2, respectivamente. Para um dado sistema, somente um dos coeficientes de erro estáticos é finito e significativo. (Quanto maior o valor do coeficiente de erro estático finito, maior será o ganho de malha conforme ω tende a zero.)

O tipo do sistema determina a inclinação da curva do log-módulo em baixas freqüências. Portanto, a informação relativa à existência e amplitude do erro em regime estacionário de um sistema de controle, para uma dada entrada, pode ser determinada a partir da observação da região de baixa freqüência na curva do log-módulo.

Determinação do coeficiente de erro de posição estático. A Fig. 9.18 mostra um exemplo do gráfico do log-módulo para um sistema tipo 0. Nos sistemas deste tipo, o módulo de $G(j\omega)\ H(j\omega)$ é igual a K_p em baixas freqüências, ou

$$\lim_{\omega\to 0} G(j\omega)H(j\omega) = K_p$$

Segue-se que a assíntota em baixa-freqüência é uma reta horizontal em $20 \log K_p$ db.

Determinação do coeficiente de erro de velocidade estático. A Fig. 9.19 mostra um exemplo do gráfico do log-módulo de um sistema tipo 1. A intersecção do segmento inicial −20-db/década (ou seu prolongamento) com a reta $\omega = 1$ tem a

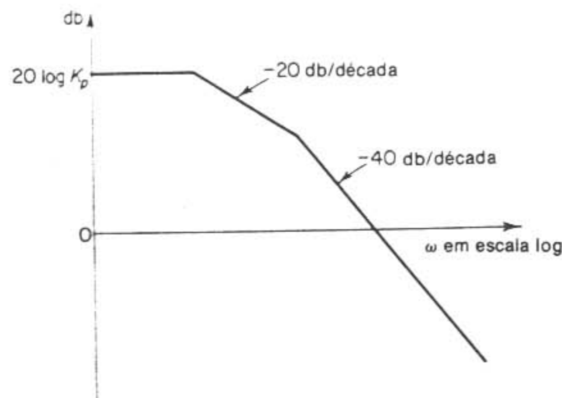


Fig. 9.18 Curva log-módulo de um sistema tipo 0.

ordenada $20 \log K_v$. Este resultado pode ser verificado da seguinte maneira: Em um sistema tipo 1,

$$G(j\omega)H(j\omega) = \frac{K_v}{j\omega} \qquad \text{para } \omega \ll 1$$

Portanto,

$$20 \log \left|\frac{K_v}{j\omega}\right|_{\omega=1} = 20 \log K_v$$

A intersecção do segmento inicial −20-db/década (ou seu prolongamento) com a reta de 0-db possui uma freqüência numericamente igual a K_v. Para verificar este resultado, vamos definir a freqüência nesta intersecção como sendo ω_1; então,

$$\left|\frac{K_v}{j\omega_1}\right| = 1$$

ou

$$K_v = \omega_1$$

Por exemplo, considere o sistema tipo 1 com realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{K}{s(Js + F)}$$

Se definirmos a freqüência de canto como ω_2 e a freqüência na intersecção do segmento −40-db/década (ou seu prolongamento) com a reta de 0-db como ω_3, então

$$\omega_2 = \frac{F}{J}, \qquad \omega_3^2 = \frac{K}{J}$$

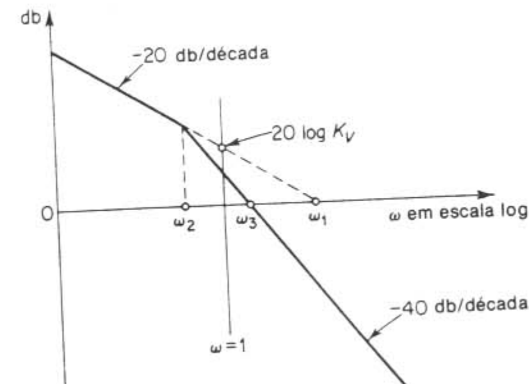


Fig. 9.19 Curva log-módulo de um sistema tipo 1.

Desde que

$$\omega_1 = K_v = \frac{K}{F}$$

segue que

$$\omega_1 \omega_2 = \omega_3^2$$

ou

$$\frac{\omega_1}{\omega_3} = \frac{\omega_3}{\omega_2}$$

No gráfico logarítmico,

$$\log \omega_1 - \log \omega_3 = \log \omega_3 - \log \omega_2$$

Portanto, o ponto ω_3 é o ponto médio entre os pontos ω_2 e ω_1. A relação de amortecimento ζ do sistema é, então,

$$\zeta = \frac{F}{2\sqrt{KJ}} = \frac{\omega_2}{2\omega_3}$$

Determinação do coeficiente de erro de aceleração estático. A Fig. 9.20 mostra um exemplo do gráfico do log-módulo para um sistema tipo 2. A intersecção do segmento inicial −40-db/década (ou seu prolongamento) com a reta $\omega = 1$ possui a ordenada 20 log K_a. Desde que em baixas freqüências

$$G(j\omega)H(j\omega) = \frac{K_a}{(j\omega)^2}$$

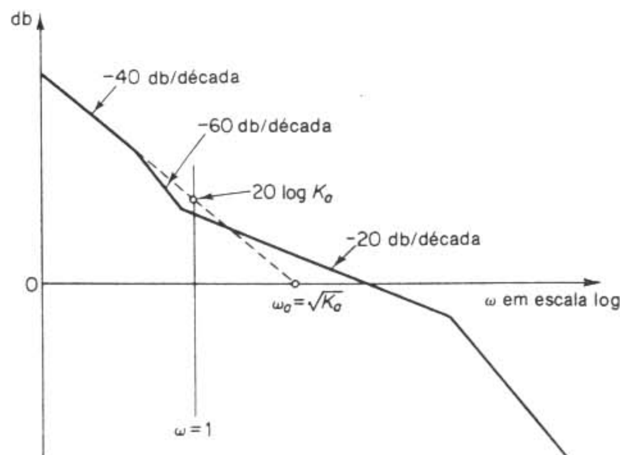


Fig. 9.20 Curva do log-módulo de um sistema tipo 2.

segue que

$$20 \log \left| \frac{K_a}{(j\omega)^2} \right|_{\omega=1} = 20 \log K_a$$

A freqüência ω_a na intersecção do segmento inicial −40-db/década (ou seu prolongamento) com a reta de 0-db fornece a raiz quadrada de K_a numericamente. Esta afirmação pode ser verificada a partir do seguinte:

$$20 \log \left| \frac{K_a}{(j\omega_a)^2} \right| = 20 \log 1 = 0$$

que fornece

$$\omega_a = \sqrt{K_a}$$

9.3 GRÁFICOS POLARES

O gráfico polar de uma função de transferência senoidal $G(j\omega)$ é um gráfico do módulo de $G(j\omega)$ *versus* o ângulo de fase de $G(j\omega)$ em coordenadas polares, conforme ω é variado desde zero até infinito. Portanto, o gráfico polar é o lugar dos vetores $|G(j\omega)| \underline{/G(j\omega)}$ conforme ω varia de zero a infinito. Observe que em gráficos polares, um ângulo de fase positivo (negativo) é medido no sentido anti-horário (horário) em relação ao eixo real positivo. O gráfico polar é muitas vezes denominado gráfico de Nyquist. Um exemplo deste tipo de gráfico é indicado na Fig. 9.21. Cada ponto no gráfico polar de $G(j\omega)$ representa o ponto terminal de um vetor, para um particular valor de ω. No gráfico polar, é importante indicar a graduação em freqüência do lugar geométrico. As projeções de $G(j\omega)$ sobre os eixos real e imaginário correspondem às suas componentes real e imaginária. Tanto o módulo $|G(j\omega)|$ como o ângulo de fase $\underline{/G(j\omega)}$ devem ser calculados diretamente para cada freqüência ω na construção dos gráficos polares. Como o gráfico logarítmico é fácil de construir, entretanto, os dados necessários para a construção do gráfico polar

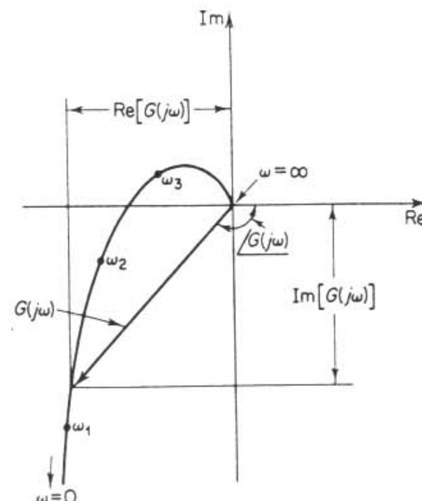


Fig. 9.21 Gráfico polar.

podem ser obtidos diretamente do gráfico logarítmico, se este último for desenhado inicialmente e os decibéis forem convertidos em módulos comuns pelo uso da Fig. 9.6.

Para dois sistemas ligados em cascata, a função de transferência global da combinação, na ausência de efeitos de carga, é o produto das duas funções de transferência individuais. Se for necessária a multiplicação de duas funções de transferência senoidais, isto pode ser realizado multiplicando-se, para cada freqüência, as funções de transferências senoidais individuais por meio da multiplicação de álgebra-complexa. Isto é, se $G(j\omega) = G_1(j\omega)G_2(j\omega)$, então

$$G(j\omega) = |G(j\omega)| \underline{/G(j\omega)}$$

onde

$$|G(j\omega)| = |G_1(j\omega)| \cdot |G_2(j\omega)|$$

e

$$\underline{/G(j\omega)} = \underline{/G_1(j\omega)} + \underline{/G_2(j\omega)}$$

O produto de $G_1(j\omega)$ e $G_2(j\omega)$ é indicado na Fig. 9.22.

Em geral, se for desejado um gráfico polar de $G_1(j\omega)G_2(j\omega)$, é conveniente desenhar primeiro um gráfico logarítmico de $G_1(j\omega)G_2(j\omega)$ e então converter em um gráfico polar ao invés de desenhar os gráficos polares de $G_1(j\omega)$ e $G_2(j\omega)$ e multiplicar estes dois no plano complexo para obter um gráfico polar de $G_1(j\omega)G_2(j\omega)$.

Uma vantagem no uso do gráfico polar é que ele mostra as características de resposta em freqüência de um sistema, em toda a faixa de freqüência, em um único gráfico. Uma desvantagem é que o gráfico não indica claramente as contribuições de cada um dos fatores individuais da função de transferência de malha-aberta.

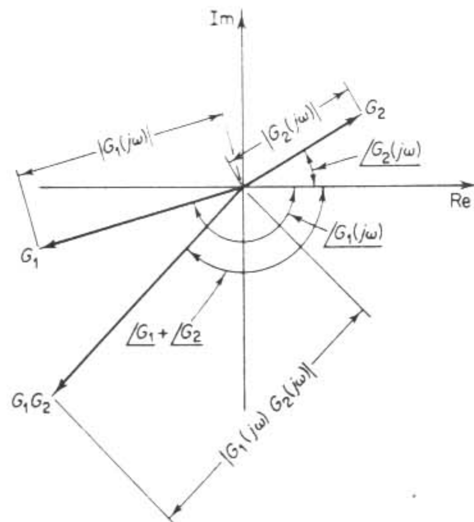


Fig. 9.22 Gráficos polares de $G_1(j\omega)$, $G_2(j\omega)$ e $G_1(j\omega)G_2(j\omega)$.

Fatores integral e derivativo $(j\omega)^{\mp 1}$. O gráfico polar de $G(j\omega) = 1/j\omega$ é o eixo imaginário negativo desde que

$$G(j\omega) = \frac{1}{j\omega} = -j\frac{1}{\omega} = \frac{1}{\omega}\underline{/-90^\circ}$$

O gráfico polar de $G(j\omega) = j\omega$ é o eixo imaginário positivo.

Fatores de primeira-ordem $(1 + j\omega T)^{\mp 1}$. Para a função de transferência senoidal

$$G(j\omega) = \frac{1}{1 + j\omega T} = \frac{1}{\sqrt{1 + \omega^2 T^2}}\underline{/-\operatorname{tg}^{-1}\omega T}$$

os valores de $G(j\omega)$ em $\omega = 0$ e $\omega = 1/T$ são, respectivamente,

$$G(j0) = 1\underline{/0^\circ} \qquad \text{e} \qquad G\left(j\frac{1}{T}\right) = \frac{1}{\sqrt{2}}\underline{/-45^\circ}$$

Se ω tende a infinito, o módulo de $G(j\omega)$ tende a zero e o ângulo de fase tende a -90°. O gráfico polar desta função de transferência é uma semicircunferência* conforme a freqüência ω varia desde zero até infinito, conforme a Fig. 9.23(a). O centro está localizado no ponto 0,5 sobre o eixo real e o raio é igual a 0,5.

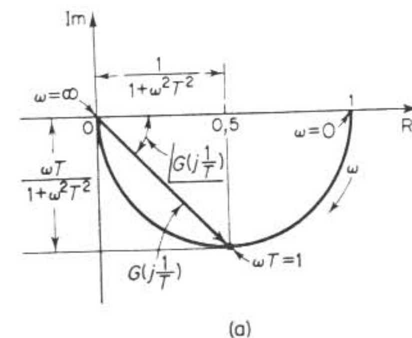
Para demonstrar que o gráfico polar é uma semicircunferência, definimos

$$G(j\omega) = X + jY$$

onde

$$X = \frac{1}{1 + \omega^2 T^2} = \text{parte real de } G(j\omega)$$

$$Y = \frac{-\omega T}{1 + \omega^2 T^2} = \text{parte imaginária de } G(j\omega)$$



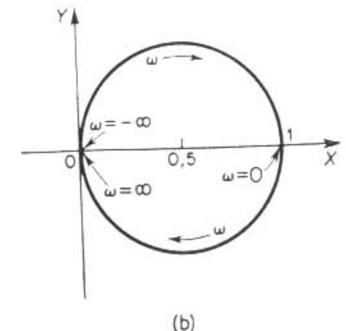


Fig. 9.23 (a) Gráfico polar de $1/(1 + j\omega T)$; (b) gráfico de $G(j\omega)$ no plano X-Y.

*N. do T.: No texto original em inglês, o autor refere-se a este lugar geométrico como um círculo em todo o texto, porém, criteriosamente, deveria considerar apenas a circunferência. Conseqüentemente, utilizaremos nesta tradução a designação deste lugar geométrico por circunferência.

Então, obtemos

$$\left(X - \frac{1}{2}\right)^2 + Y^2 = \left(\frac{1}{2}\frac{1 - \omega^2 T^2}{1 + \omega^2 T^2}\right)^2 + \left(\frac{-\omega T}{1 + \omega^2 T^2}\right)^2 = \left(\frac{1}{2}\right)^2$$

Portanto, no plano X-Y, $G(j\omega)$ é uma circunferência com centro em $X = 1/2$, $Y = 0$ e com raio 1/2, conforme indicado na Fig. 9.23(b). A semicircunferência inferior corresponde a $0 \leq \omega \leq \infty$ e a semicircunferência superior corresponde a $-\infty \leq \omega \leq 0$.

O gráfico polar da função de transferência $1 + j\omega T$ é simplesmente a metade superior da reta passando pelo ponto (1, 0) no plano complexo e paralela ao eixo imaginário conforme a Fig. 9.24.* O gráfico polar de $1 + j\omega T$ possui uma aparência completamente diferente daquela de $1/(1 + j\omega T)$.

Fatores quadráticos $[1 + 2\zeta(j\omega/\omega_n) + (j\omega/\omega_n)^2]^{\mp 1}$. As partes de baixa-freqüência e alta-freqüência do gráfico polar para a seguinte função de transferência senoidal:

$$G(j\omega) = \frac{1}{1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2} \qquad (\zeta > 0)$$

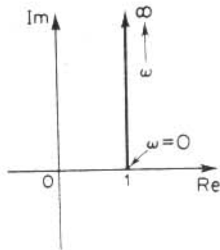


Fig. 9.24 Gráfico polar de $1 + j\omega T$.

são dadas, respectivamente, por

$$\lim_{\omega \to 0} G(j\omega) = 1\underline{/0^\circ} \qquad \text{e} \qquad \lim_{\omega \to \infty} G(j\omega) = 0\underline{/-180^\circ}$$

O gráfico polar desta função de transferência senoidal tem início em $1\underline{/0^\circ}$ e termina em $0\underline{/-180^\circ}$ conforme ω aumenta desde zero até infinito. Portanto, a parte de alta-freqüência de $G(j\omega)$ é tangente ao eixo real negativo. Os valores de $G(j\omega)$ no intervalo de freqüência de interesse podem ser calculados diretamente ou usando-se o gráfico logarítmico.

Exemplos de gráficos polares da função de transferência considerada são indicados na Fig. 9.25. A forma exata de um gráfico polar depende do valor da relação de amortecimento ζ, porém a forma geral do gráfico é a mesma tanto para o caso subamortecido ($1 > \zeta > 0$) como para o caso superamortecido ($\zeta > 1$).

*N. do T. O autor está considerando apenas $\omega \geq 0$.

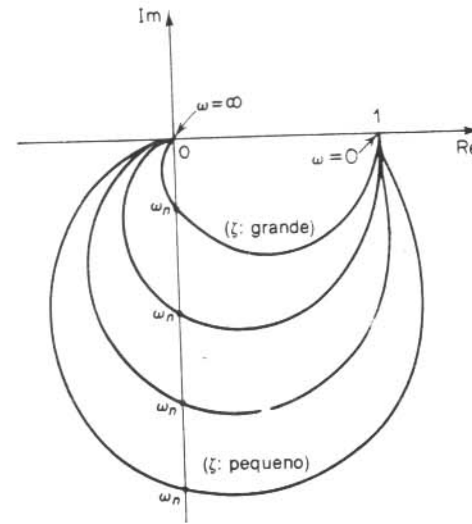


Fig. 9.25 Gráficos polares de $\dfrac{1}{1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2}$, $(\zeta > 0)$.

Para o caso subamortecido em $\omega = \omega_n$, temos que $G(j\omega_n) = 1/(j2\zeta)$ e o ângulo de fase em $\omega = \omega_n$ é -90°. Além disso, pode ser observado que a freqüência na qual o lugar geométrico de $G(j\omega)$ intercepta o eixo imaginário é a freqüência natural não amortecida, ω_n. No gráfico polar, o ponto de freqüência cuja distância à origem é máxima corresponde à freqüência de ressonância ω_r. O valor de pico de $G(j\omega)$ é obtido pela relação entre o módulo do vetor na freqüência de ressonância ω_r e o módulo do vetor em $\omega = 0$. A freqüência de ressonância ω_r é indicada no gráfico polar mostrado na Fig. 9.26.

Para o caso superamortecido, conforme ζ aumenta bem além da unidade, o lugar geométrico de $G(j\omega)$ aproxima-se de uma semicircunferência. Isto pode ser observado pelo fato de que para um sistema muito amortecido as raízes características são reais e uma delas é muito menor que a outra. Desde que para ζ suficientemente grande o efeito da raiz muito maior na resposta resulta muito pequeno, o sistema comporta-se como de primeira-ordem.

Para a função de transferência senoidal,

$$G(j\omega) = 1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2 = \left(1 - \frac{\omega^2}{\omega_n^2}\right) + j\left(\frac{2\zeta\omega}{\omega_n}\right)$$

a parte em baixa-freqüência da curva é

$$\lim_{\omega \to 0} G(j\omega) = 1\underline{/0^\circ}$$

e a parte em alta-freqüência é

$$\lim_{\omega \to \infty} G(j\omega) = \infty\underline{/180^\circ}$$

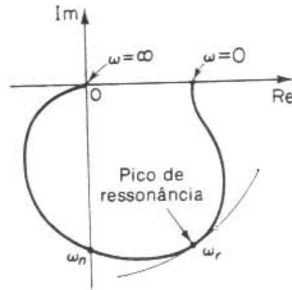


Fig. 9.26 Gráfico polar mostrando o pico de ressonância e a freqüência de ressonância ω_r.

Desde que a parte imaginária de $G(j\omega)$ é positiva para ω 0 e monotonicamente crescente e a parte real de $G(j\omega)$ é monotonicamente decrescente da unidade, a forma geral do gráfico polar de $G(j\omega)$ é conforme indicada na Fig. 9.27. O ângulo de fase está entre 0° e 180°.

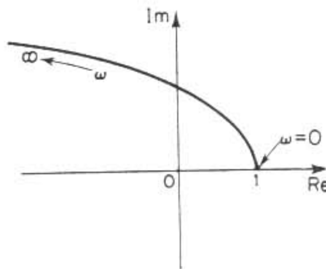


Fig. 9.27 Gráfico polar de $1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2$, $(\zeta > 0)$.

Exemplo 9.4 Considere a seguinte função de transferência de segunda-ordem:

$$G(s) = \frac{1}{s(Ts+1)}$$

Esboce um gráfico polar desta função de transferência.

Desde que a função de transferência senoidal pode ser escrita

$$G(j\omega) = \frac{1}{j\omega(1+j\omega T)}$$
$$= -\frac{T}{1+\omega^2T^2} - j\frac{1}{\omega(1+\omega^2T^2)}$$

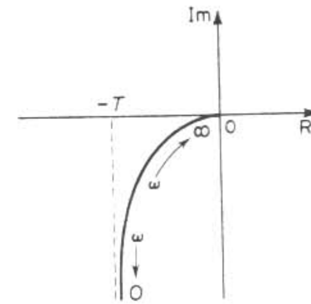


Fig. 9.28 Gráfico polar de $1/[j\omega(1+j\omega T)]$.

a parte de baixa-freqüência do gráfico polar torna-se

$$\lim_{\omega\to 0} G(j\omega) = -T - j\infty = \infty\angle{-90^\circ}$$

e a parte de alta-freqüência torna-se

$$\lim_{\omega\to\infty} G(j\omega) = 0 - j0 = 0\angle{-180^\circ}$$

A forma geral do gráfico polar de $G(j\omega)$ é indicada na Fig. 9.28. O gráfico de $G(j\omega)$ é assintótico para a reta vertical passando pelo ponto $(-T, 0)$. Desde que esta função de transferência envolve a integração $(1/s)$, a forma geral do gráfico polar difere substancialmente daquelas correspondentes a funções de transferência de segunda-ordem que não possuem a integração.

Atraso de transporte. O atraso de transporte

$$G(j\omega) = e^{-j\omega T}$$

pode ser escrito

$$G(j\omega) = 1\angle{\cos\omega T - j\,\text{sen}\,\omega T}$$

Desde que o módulo de $G(j\omega)$ é sempre unitário e o ângulo de fase varia linearmente com ω, o gráfico polar do atraso de transporte é uma circunferência unitária, conforme indicada na Fig. 9.29.

Em baixas freqüências, o atraso de transporte $e^{-j\omega T}$ e o atraso de primeira-ordem $1/(1+j\omega T)$ se comportam similarmente, conforme é mostrado na Fig. 9.30. Os gráficos polares de $e^{-j\omega T}$ e $1/(1+j\omega T)$ são tangentes entre si na freqüência $\omega = 0$. Este resultado pode ser comprovado pelo fato de que, para $\omega \ll 1/T$,

$$e^{-j\omega T} \doteq 1 - j\omega T \qquad \text{e} \qquad \frac{1}{1+j\omega T} \doteq 1 - j\omega T$$

Para $\omega \gg 1/T$, entretanto, existe uma diferença essencial entre $e^{-j\omega T}$ e $1/(1+j\omega T)$ como pode ser visto na Fig. 9.30.

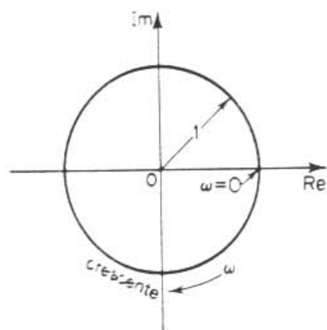
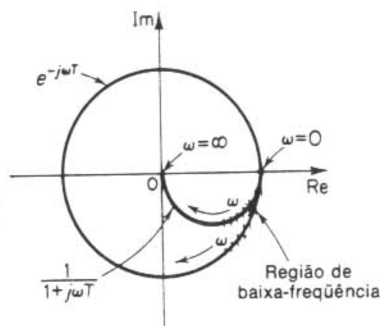


Fig. 9.29 Gráfico polar do atraso de transporte.

Fig. 9.30 Gráficos polares de $e^{-j\omega T}$ e $1/(1 + j\omega T)$.



Exempl[illegible] Obtenha o gráfico polar da seguinte função de transferência:

$$G(j\omega) = \frac{e^{-j\omega L}}{1 + j\omega T}$$

Desde que $G(j\omega)$ pode ser escrita

$$G(j\omega) = (e^{-j\omega L})\left(\frac{1}{1 + j\omega T}\right)$$

o módulo e o ângulo de fase são, respectivamente,

$$|G(j\omega)| = |e^{-j\omega L}| \cdot \left|\frac{1}{1 + j\omega T}\right| = \frac{1}{\sqrt{1 + \omega^2 T^2}}$$

e

$$\underline{/G(j\omega)} = \underline{/e^{-j\omega L}} + \underline{/\frac{1}{1 + j\omega T}} = -\omega L - \text{tg}^{-1}\ \omega T$$

Desde que o módulo diminui monotonicamente desde a unidade e o ângulo de fase também diminui monotônica e indefinidamente, o gráfico polar da função de transferência em estudo é uma espiral, conforme indicado na Fig. 9.31.

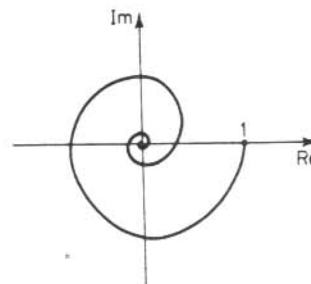


Fig. 9.31 Gráfico polar de $e^{-j\omega L}/(1 + j\omega T)$.

Formas gerais de gráficos polares. Os gráficos polares de uma função de transferência da forma

$$G(j\omega) = \frac{K(1 + j\omega T_a)(1 + j\omega T_b)\cdots}{(j\omega)^\lambda(1 + j\omega T_1)(1 + j\omega T_2)\cdots}$$
$$= \frac{b_0(j\omega)^m + b_1(j\omega)^{m-1} + \cdots}{a_0(j\omega)^n + a_1(j\omega)^{n-1} + \cdots}$$

onde o grau do polinômio do denominador é maior do que o grau do polinômio do numerador, apresentarão as seguintes formas gerais:

1. *Para* $\lambda = 0$ *ou sistemas tipo* 0: O ponto inicial do gráfico polar (que corresponde a $\omega = 0$.) é finito e está sobre o eixo real positivo. A tangente do gráfico polar em $\omega = 0$ é perpendicular ao eixo real. O ponto final, que corresponde a $\omega = \infty$, corresponde à origem, e a curva é tangente a um dos eixos.
2. *Para* $\lambda = 1$ *ou sistemas tipo* 1: O termo $j\omega$ no denominador contribui com $-90°$ para o ângulo de fase total de $G(j\omega)$ para $0 \leq \omega \leq \infty$. Em $\omega = 0$, o módulo de $G(j\omega)$ é infinito, e o ângulo de fase resulta $-90°$. Em baixas freqüências, o gráfico polar é assintótico a uma reta paralela ao eixo imaginário negativo. Em $\omega = \infty$, o módulo torna-se nulo e a curva converge para a origem e é tangente a um dos eixos.
3. *Para* $\lambda = 2$ *ou sistemas do tipo 2:* O termo $(j\omega)^2$ no denominador contribui com $-180°$ para o ângulo de fase total de $G(j\omega)$ para $0 \leq \omega \leq \infty$. Em $\omega = 0$, o

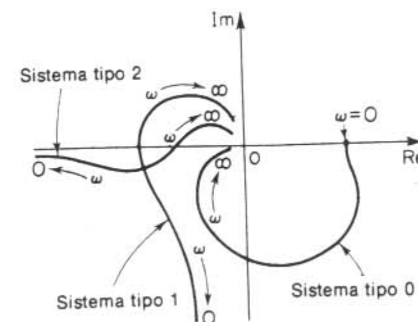
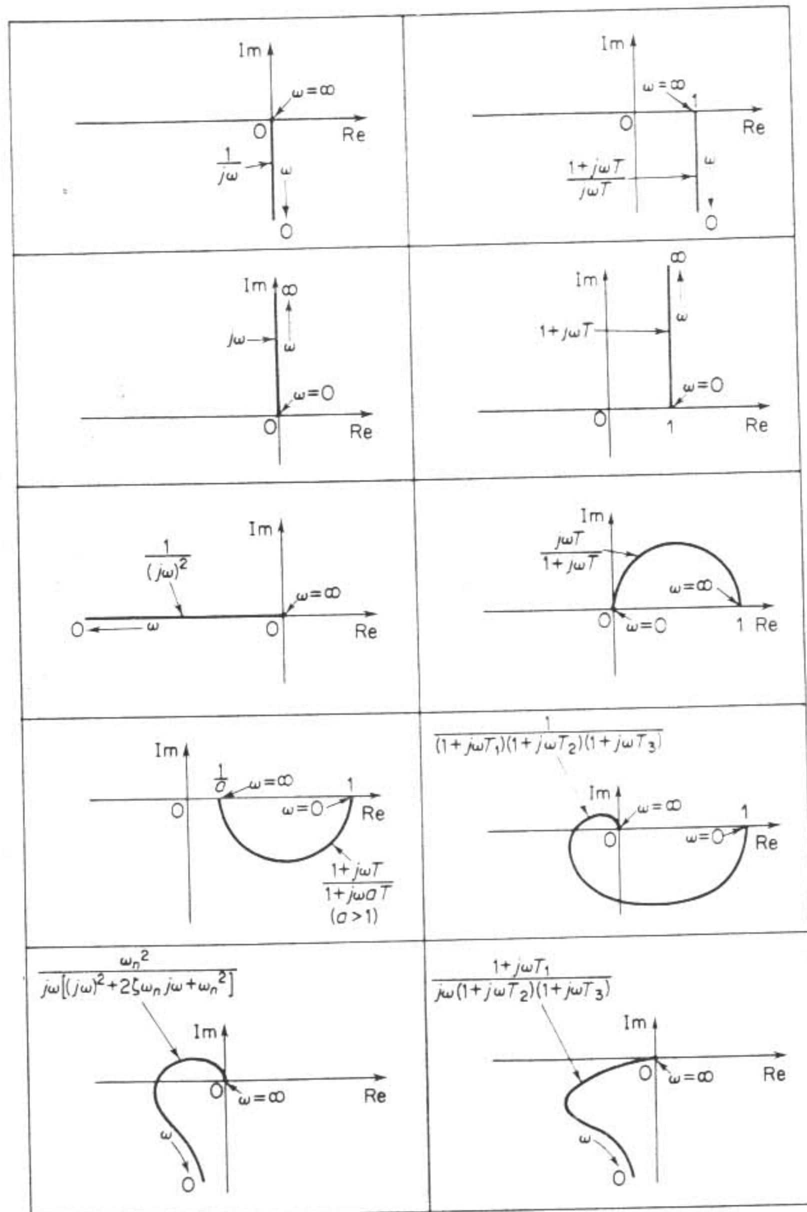


Fig. 9.32 Gráficos polares de sistemas tipo 0, tipo 1 e tipo 2.

Tabela 9.1 Gráficos polares de funções de transferência simples



módulo de $G(j\omega)$ é infinito, e o ângulo de fase é igual a $-180°$. Em baixas freqüências, o gráfico polar é assintótico a uma reta paralela ao eixo real negativo. Em $\omega = \infty$, o módulo torna-se nulo e a curva é tangente a um dos eixos.

Os formatos gerais das partes de baixa-freqüência dos gráficos polares de sistemas tipo 0, tipo 1 e tipo 2 são indicados na Fig. 9.32. Pode ser visto que, se o grau do polinômio do denominador de $G(j\omega)$ é maior do que o do numerador, então os lugares geométricos de $G(j\omega)$ convergem para a origem no sentido horário. Em $\omega = \infty$, os lugares geométricos são tangentes a um ou ao outro eixo, conforme indicado na Fig. 9.33.

Para o caso em que os graus dos polinômios do denominador e do numerador de $G(j\omega)$ são os mesmos, o gráfico polar tem início a uma distância finita sobre o eixo real e termina em um ponto finito sobre o eixo real.

Note que quaisquer formas complicadas nas curvas do gráfico polar são causadas pela dinâmica do numerador, isto é, pelas constantes de tempo no nume-

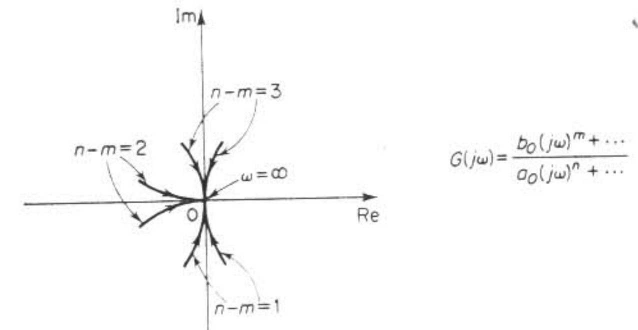


Fig. 9.33 Gráficos polares na faixa de alta-freqüência.

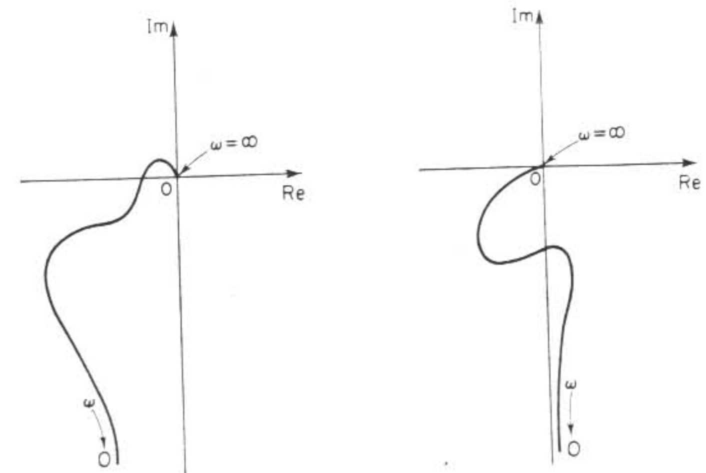


Fig. 9.34 Gráficos polares de funções de transferência com dinâmica no numerador.

rador da função de transferência. A Fig. 9.34 fornece exemplos de gráficos polares de funções de transferência com dinâmicas no numerador. Na análise de sistemas de controle, o gráfico polar de $G(j\omega)$ na faixa de freqüência de interesse deve ser precisamente determinado.

A Tabela 9.1 mostra esboços de gráficos polares para algumas funções de transferência.

9.4 GRÁFICOS DO LOG-MÓDULO *VERSUS* FASE

Uma outra abordagem para retratar graficamente as características de resposta em freqüência é usar o gráfico do log-módulo *versus* fase, que é um gráfico do logaritmo do módulo em decibéis *versus* o ângulo de fase, ou margem de fase, para uma faixa de freqüência de interesse. (A margem de fase é a diferença entre o ângulo de fase real ϕ e $-180°$; isto é, $\phi - (-180°) = 180° + \phi$.) A curva é graduada em termos da freqüência ω. Estes gráficos do log-módulo *versus* fase são algumas vezes denominados gráficos de Nichols.

No diagrama de Bode, as características de resposta em freqüência de $G(j\omega)$ são fornecidas em um papel monolog por meio de duas curvas separadas, a curva do log-módulo e a curva do ângulo de fase, enquanto no gráfico do log-módulo *versus* fase, as duas curvas do diagrama de Bode são combinadas em apenas uma. O gráfico do log-módulo *versus* fase pode ser facilmente construído pelos valores obtidos do logaritmo do módulo e do ângulo de fase a partir do diagrama de Bode. Note que no gráfico do log-módulo *versus* fase, uma variação na constante de ganho de $G(j\omega)$ simplesmente desloca a curva para cima (para ganhos crescentes) ou para baixo (para ganho decrescente), porém a forma da curva permanece a mesma.

As vantagens do gráfico do log-módulo *versus* fase são que a estabilidade relativa do sistema de malha-fechada pode ser determinada rapidamente e que a compensação pode ser realizada com facilidade.

Os gráficos do log-módulo *versus* fase para as funções de transferência senoidais $G(j\omega)$ e $1/G(j\omega)$ são anti-simétricos em relação à origem desde que

$$\left|\frac{1}{G(j\omega)}\right| \text{em db} = -|G(j\omega)| \text{ em db}$$

e

$$\angle\frac{1}{G(j\omega)} = -\angle G(j\omega)$$

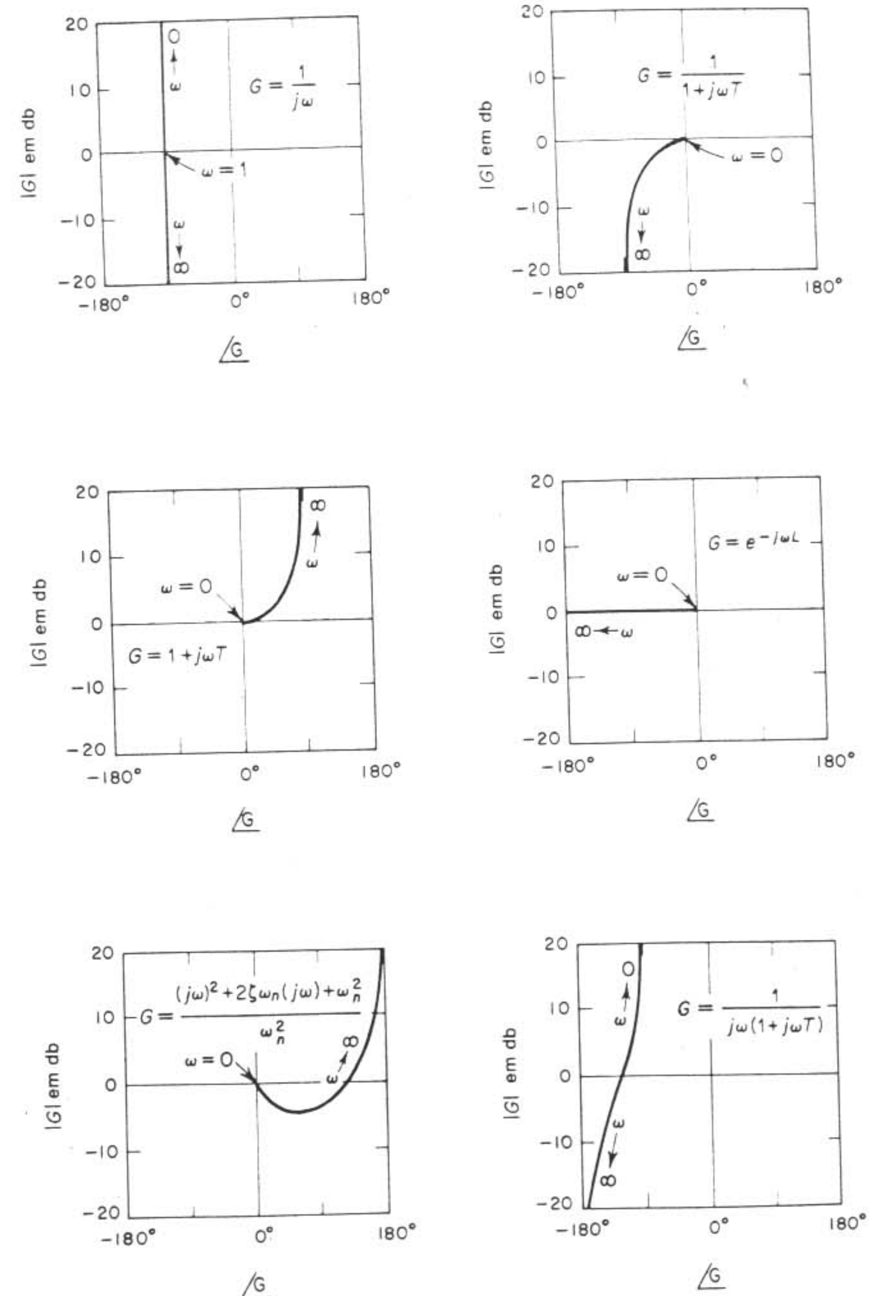
Desde que as características do log-módulo e do ângulo de fase das funções de transferência básicas foram discutidas com detalhes nas Seções 9.2 e 9.3, aqui será suficiente fornecer exemplos de alguns gráficos do log-módulo *versus* fase. A Tabela 9.2 mostra estes exemplos.

A Fig. 9.35 compara as curvas de resposta em freqüência de

$$G(j\omega) = \frac{1}{1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2}$$

nas três representações diferentes. No gráfico log-módulo *versus* fase, a distância vertical entre os pontos correspondentes a $\omega = 0$ e $\omega = \omega_r$, onde ω_r é a freqüência de ressonância, é o valor de pico de $G(j\omega)$ em decibéis.

Tabela 9.2 Gráficos log-módulo *versus* fase de funções de transferência simples



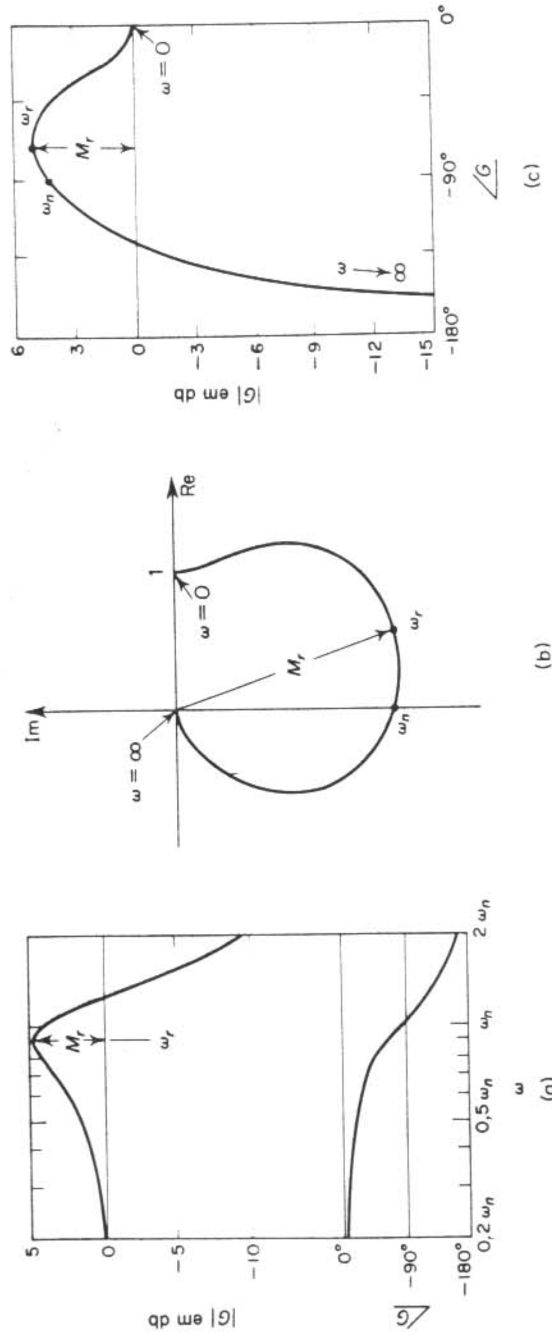


Fig. 9.35 Três representações da resposta em freqüência de

$$\frac{1}{1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2}, \qquad (\zeta > 0).$$

(a) Gráfico logarítmico; (b) gráfico polar; (c) gráfico log-módulo *versus* fase.

9.5 CRITÉRIO DE ESTABILIDADE DE NYQUIST

Esta seção apresenta o critério de estabilidade de Nyquist e a base matemática associada.

Considere o sistema de malha-fechada indicado na Fig. 9.36. A função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)}$$

Para estabilidade, todas as raízes da equação característica

$$1 + G(s)H(s) = 0$$

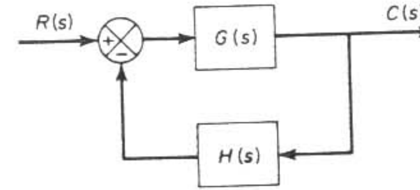


Fig. 9.36 Sistema de malha-fechada.

devem estar no semiplano esquerdo do plano s. O critério de estabilidade de Nyquist relaciona a resposta em freqüência de malha-aberta $G(j\omega)H(j\omega)$ ao número de zeros e pólos de $1 + G(s)H(s)$ que estão no semiplano direito do plano s. Este critério, devido a H. Nyquist, é útil em engenharia de controle porque a estabilidade absoluta do sistema de malha-fechada pode ser determinada graficamente a partir das curvas de resposta em freqüência de malha-aberta e não há necessidade para determinar realmente os pólos de malha-fechada. As curvas de resposta em freqüência de malha-aberta obtidas analiticamente, bem como aquelas obtidas experimentalmente, podem ser utilizadas na análise de estabilidade. Isto é conveniente porque, no projeto de um sistema de controle, muitas vezes ocorre que as expressões matemáticas para alguns dos componentes não são conhecidas, sendo disponíveis apenas seus dados de resposta em freqüência.

O critério de estabilidade de Nyquist é baseado em um teorema da teoria de variáveis complexas. Para entender o critério, discutiremos inicialmente mapeamentos de contornos no plano complexo.

Suporemos que a função de transferência de malha-aberta $G(s)H(s)$ é representável como uma relação de polinômios em s. Para um sistema fisicamente realizável, o grau do polinômio do denominador da função de transferência de malha-fechada deve ser maior ou igual àquele do polinômio do numerador. Isto significa que o limite de $G(s)H(s)$ é nulo ou constante, para s tendendo a infinito, para qualquer sistema fisicamente realizável.

Estudo preliminar. A equação característica do sistema indicado na Fig. 9.36 é

$$F(s) = 1 + G(s)H(s) = 0$$

Mostraremos que para um dado caminho fechado e contínuo no plano s, que nã passa através de quaisquer pontos singulares, corresponde uma curva fechada n

plano $F(s)$. O número e sentido de envolvimento da origem do plano $F(s)$ pela curva fechada representam um papel particularmente importante no que se segue, para mais tarde correlacionarmos o número e sentidos de envolvimentos à estabilidade do sistema.

Considere, por exemplo, a seguinte função de transferência de malha-aberta:

$$G(s)H(s) = \frac{6}{(s+1)(s+2)}$$

A equação característica é

$$F(s) = 1 + G(s)H(s) = 1 + \frac{6}{(s+1)(s+2)}$$

$$= \frac{(s+1{,}5+j2{,}4)(s+1{,}5-j2{,}4)}{(s+1)(s+2)} = 0$$

A função $F(s)$ é analítica em todos os pontos do plano s, exceto em seus pontos singulares. Para cada ponto de analiticidade no plano s, corresponderá um ponto no plano $F(s)$. Por exemplo, se $s = 1 + j2$, então $F(s)$ torna-se

$$F(1+j2) = 1 + \frac{6}{(2+j2)(3+j2)}$$

$$= 1{,}12 - j0{,}58$$

Portanto, o ponto $s = 1 + j2$ no plano s mapeia no ponto $1{,}12 - j0{,}58$ no plano $F(s)$.

Portanto, conforme anteriormente estabelecido, para um dado caminho fechado contínuo no plano s, que não passa por pontos singulares, corresponderá uma curva fechada no plano $F(s)$. A Fig. 9.37(a) mostra mapeamentos conformes das retas $\omega = 0, 1, 2, 3$ e das retas $\sigma = 1, 0, -1, -2, -3, -4$ no semiplano superior do plano s, para o plano $F(s)$. Por exemplo, a reta $s = j\omega$ no semiplano superior do plano s ($\omega \geq 0$) mapeia na curva indicada por $\sigma = 0$ no plano $F(s)$. A Fig. 9.37(b) mostra mapeamentos conformes das retas $\omega = 0, -1, -2, -3$ e das retas $\sigma = 1, 0, -1, -2, -3, -4$ do semiplano inferior do plano s, no plano $F(s)$. Observe que para um dado σ, a curva para freqüências negativas é simétrica em relação ao eixo real com a curva para freqüências positivas. Referindo-se às Figs. 9.37(a) e (b), verificamos que para o caminho $ABCD$ no plano s percorrido no sentido horário, a curva correspondente no plano $F(s)$ é $A'\ B'\ C'$ e D'. As setas sobre as curvas indicam sentidos de percurso. Analogamente, o caminho $DEFA$ no plano s mapeia na curva D', E', F', A', no plano $F(s)$. Devido à propriedade de mapeamento conforme, os ângulos correspondentes no plano s e no plano $F(s)$ são iguais e possuem o mesmo sentido. (Por exemplo, desde que as retas AB e BC se interceptam em ângulos retos entre si no plano s, as curvas $A'B'$ e $B'C'$ também se interceptam em ângulos retos no ponto B' do plano $F(s)$. Referindo-se à Fig. 9.37(c), verificamos que sobre o contorno fechado $ABCDEFA$ no plano s, a variável s tem início no ponto A e assume valores neste caminho em um sentido horário até retornar ao ponto de partida A. A curva correspondente no plano $F(s)$ é denotada por $A'B'C'D'E'F'A'$. Se definirmos a área à direita do contorno, quando um ponto representativo s move-se no sentido horário, como correspondente ao lado de dentro do contorno e a área à esquerda como o lado de fora, então a área hachurada na Fig. 9.37(c) está envolvida pelo contorno $ABCDEFA$ e está dentro dele. Da Fig. 9.37(c) pode-se verificar que,

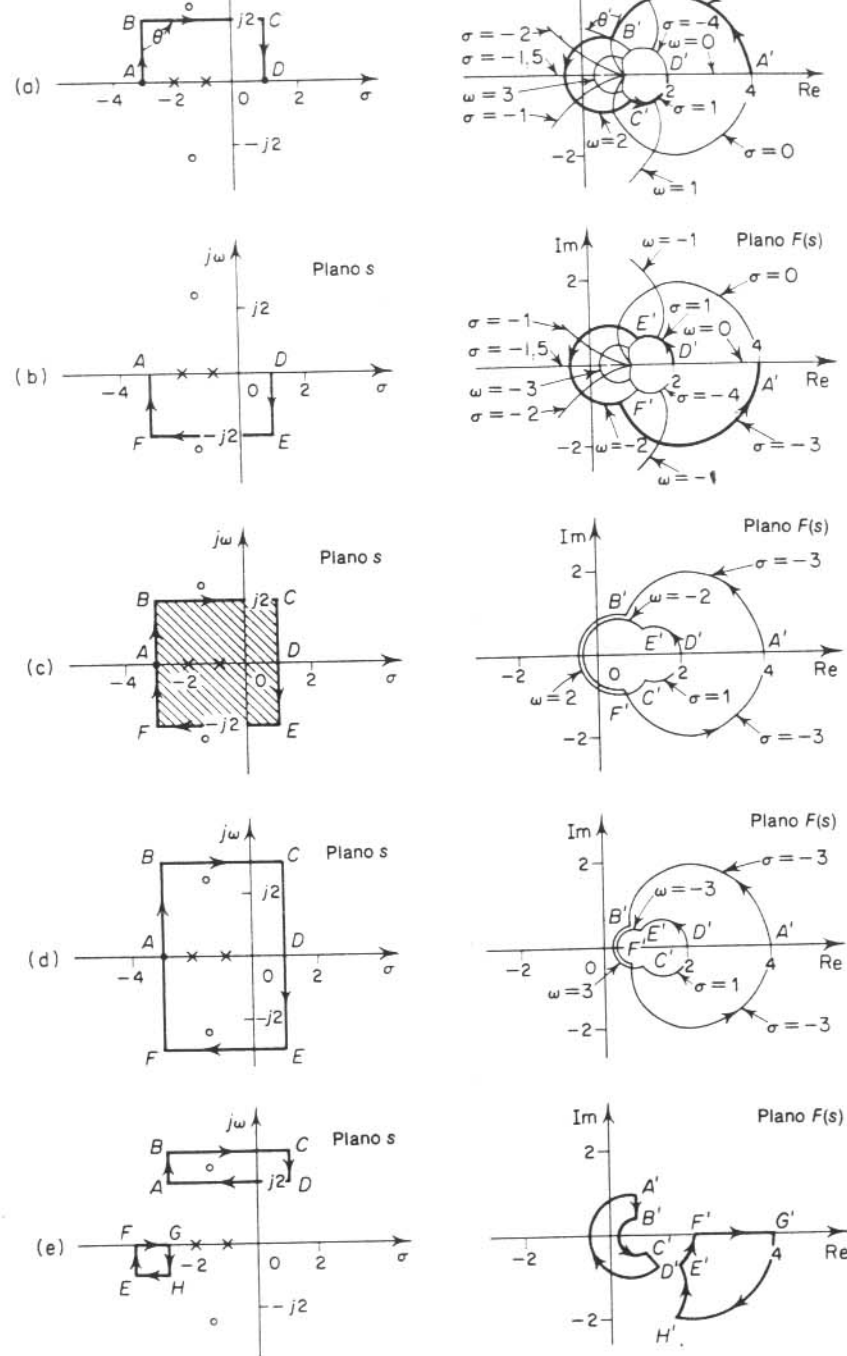


Fig. 9.37 Mapeamento conforme de grades no plano s para o plano $F(s)$.

quando o contorno no plano s envolve dois pólos de $F(s)$, o lugar geométrico de $F(s)$ envolve a origem do plano $F(s)$ duas vezes, no sentido anti-horário.

O número de envolvimentos da origem do plano $F(s)$ depende do contorno fechado no plano s. Se este contorno envolve dois zeros e dois pólos de $F(s)$, então o lugar geométrico de $F(s)$ correspondente não envolve a origem, conforme indicado na Fig. 9.37(d). Se este contorno envolve apenas um zero, o lugar geométrico correspondente de $F(s)$ envolve a origem uma vez no sentido horário. Isto é mostrado na Fig. 9. 37(e). Finalmente, se o contorno fechado no plano s não envolve nem zeros nem pólos, então o lugar geométrico de $F(s)$ não envolve a origem do plano $F(s)$. Este resultado também é mostrado na Fig. 9.37(e).

Note que para cada ponto no plano s, exceto para os pontos singulares, há apenas um ponto correspondente no plano $F(s)$; isto é, o mapeamento do plano s no plano $F(s)$ é um-a-um. O mapeamento do plano $F(s)$ no plano s pode não ser um-a-um, entretanto, de modo que um dado ponto no plano $F(s)$ pode corresponder a mais do que um ponto no plano s. Por exemplo, o ponto B' no plano $F(s)$ na Fig. 9.37(d) corresponde aos dois pontos $(-3; 3)$ e $(0; -3)$ no plano s.

Da análise anterior, podemos observar que o sentido de envolvimento da origem do plano $F(s)$ depende de o contorno no plano s envolver um pólo ou um zero. Note que a localização de um pólo ou zero no plano s, ou no semiplano direito ou no semiplano esquerdo do plano s, não faz qualquer diferença, porém o envolvimento de um pólo ou de um zero faz diferença. Se o contorno no plano s envolve k zeros e k pólos ($k = 0, 1, 2, \ldots$), isto é, um número igual de cada, então a curva fechada correspondente no plano $F(s)$ não envolve a origem do plano $F(s)$. A discussão anterior corresponde a uma explicação gráfica do teorema do mapeamento, que é a base para o critério de estabilidade de Nyquist.

Teorema do mapeamento. Seja $F(s)$ uma relação de dois polinômios em s. Seja P o número de pólos e Z o número de zeros de $F(s)$ que estão dentro de algum contorno fechado no plano s, levando-se em conta a multiplicidade dos pólos e zeros. Seja este contorno tal que não passe por nenhum pólo ou zero de $F(s)$. Esse contorno fechado no plano s é então mapeado no plano $F(s)$ como uma curva fechada. O número total N de envolvimentos da origem do plano $F(s)$ no sentido horário, conforme um ponto representativo s percorre todo o contorno no sentido horário, é igual a $Z - P$. (Note que por meio deste teorema do mapeamento os números de zeros e pólos não podem ser determinados, apenas sua diferença.)

Não apresentaremos aqui uma prova formal deste teorema, mas deixaremos esta prova para o Problema A.9.6. Note que o número positivo N indica um excesso de zeros em relação aos pólos da função $F(s)$ e um número negativo de N indica um excesso de pólos em relação aos zeros. Em aplicações de sistema de controle, o número P pode ser realmente determinado para $F(s) = 1 + G(s)H(s)$ a partir da função $G(s)H(s)$. Portanto, se N é determinado a partir do gráfico de $F(s)$, o número de zeros no contorno fechado do plano s pode ser realmente determinado. Note que as formas exatas do contorno no plano s e o lugar geométrico de $F(s)$ não possuem importância desde que sejam considerados os envolvimentos da origem, pois os envolvimentos dependem apenas do envolvimento de pólos e/ou zeros de $F(s)$ pelo contorno no plano s.

Aplicação do teorema do mapeamento na análise de estabilidade de sistemas de malha-fechada. Para analisar a estabilidade de sistemas de controle lineares, consideraremos o contorno fechado no plano s envolvendo todo o semiplano direito do plano s. O contorno consiste em todo o eixo $j\omega$ desde $\omega = -\infty$ até $\omega = +\infty$ e um caminho semicircular de raio infinito no semiplano direito do plano s. Este contorno é denominado caminho de Nyquist. (O sentido do caminho é o horário.) O caminho

de Nyquist envolve todo o semiplano direito do plano s e todos os pólos e zeros de $1 + G(s)H(s)$ que possuam partes reais positivas. (Se não houver zeros de $1 + G(s)H(s)$ no semiplano direito do plano s, então lá também não haverá pólos de malha-fechada, e o sistema é estável.) É necessário que o contorno fechado, ou o caminho de Nyquist, não passe através de qualquer pólo e zero de $1 + G(s)H(s)$. Se $G(s)H(s)$ possui um pólo ou pólos na origem do plano s, então o mapeamento do ponto $s = 0$ resulta indeterminado. Nestes casos, a origem é evitada, considerando-se um contorno em torno dela. (Uma discussão detalhada deste caso especial será dada mais tarde.)

Se o teorema do mapeamento for aplicado ao caso especial no qual $F(s)$ é igual a $1 + G(s)H(s)$, então podemos fazer a seguinte afirmação: Se o contorno fechado no plano s envolver todo o semiplano direito do plano s, conforme indicado na Fig. 9.38, então o número de zeros da função $F(s) = 1 + G(s)H(s)$ no semiplano direito é igual ao número de pólos da função $F(s) = 1 + G(s)H(s)$ no semiplano direito do plano s mais o número de envolvimentos da origem do plano $1 + G(s)H(s)$ no sentido horário para a curva fechada correspondente neste último plano.

Devido à condição assumida de que

$$\lim_{s\to\infty} [1 + G(s)H(s)] = \text{constante}$$

a função $1 + G(s)H(s)$ permanece constante conforme s percorre a semicircunferência de raio infinito. Por esta razão, independente de o lugar geométrico de $1 + G(s)H(s)$, o envolvimento da origem do plano $1 + G(s)H(s)$ pode ser determinado considerando-se apenas uma parte do contorno fechado no plano s, isto é, o eixo $j\omega$. Os envolvimentos da origem, se houver, ocorrerão somente enquanto um ponto representativo se move desde $-j\infty$ até $+j\infty$ ao longo do eixo $j\omega$ desde que não exista nenhum zero ou pólo no eixo $j\omega$.

Note que a parte do contorno de $1 + G(s)H(s)$ desde $\omega = -\infty$ até $\omega = \infty$ é simplesmente $1 + G(j\omega)H(j\omega)$. Desde que $1 + G(j\omega)H(j\omega)$ é a soma vetorial do vetor unitário e o vetor $G(j\omega)H(j\omega)$, $1 + G(j\omega)H(j\omega)$ é idêntico ao vetor desenhado desde o ponto $-1 + j0$ até o ponto terminal do vetor $G(j\omega)H(j\omega)$, conforme indicado na Fig. 9.39. O envolvimento da origem pelo gráfico de $1 + G(j\omega)H(j\omega)$ é equivalente ao envolvimento do ponto $-1 + j0$ pelo lugar geométrico de $G(j\omega)H(j\omega)$. Portanto, a estabilidade de um sistema de malha-fechada pode ser investigada examinando-se os envolvimentos do ponto $-1 + j0$ pelo lugar geométrico de $G(j\omega)H(j\omega)$. O número de envolvimento do ponto $-1 + j0$ no sentido horário pode ser determinado desenhando-se um vetor desde o ponto $-1 + j0$ até o lugar geométrico de $G(j\omega)H(j\omega)$, iniciando desde $\omega = -\infty$, passando por $\omega = 0$ e terminando em $\omega = +\infty$ e contando o número de rotações do vetor no sentido horário.

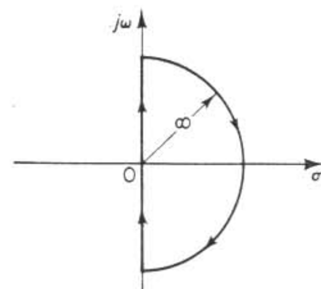


Fig. 9.38 Contorno fechado no plano s.

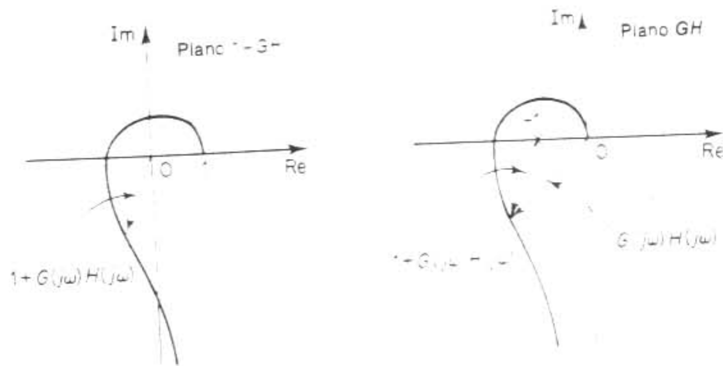


Fig. 9.39 Gráficos de $G(j\omega)H(j\omega)$ no plano $1 - GH$ e no plano GH.

A construção do gráfico de $G(j\omega)H(j\omega)$ para o caminho de Nyquist é direita. O mapa do eixo $j\omega$ negativo é a imagem especular, em relação ao eixo real, do mapa do eixo $j\omega$ positivo. Isto é, o gráfico de $G(j\omega)H(j\omega)$ e o gráfico de $G(-j\omega)H(-j\omega)$ são simétricos entre si em relação ao eixo real. A semicircunferência com raio infinito mapeia ou na origem do plano GH ou em um ponto sobre o eixo real do plano GH.

Na discussão precedente, admitiu-se que $G(s)H(s)$ é a relação de dois polinômios em s. Portanto, o atraso de transporte e^{-Ts} foi excluído da discussão. Note, entretanto, que uma discussão análoga é aplicável a sistemas com atraso de transporte, embora a prova desta asserção não seja aqui fornecida. A estabilidade de um sistema com atraso de transporte pode ser determinada a partir das curvas de resposta em freqüência de malha-aberta examinando-se o número de envolvimentos do ponto $-1 + j0$, como no caso de um sistema cuja função de transferência de malha-aberta é uma relação de dois polinômios em s.

Critério de estabilidade de Nyquist. A análise anterior, utilizando o envolvimento do ponto $-1 + j0$ pelo lugar geométrico de $G(j\omega)H(j\omega)$, é sumariada no seguinte critério de estabilidade de Nyquist:

Critério de estabilidade de Nyquist [*Para um caso especial onde $G(s)H(s)$ não possui pólos ou zeros no eixo $j\omega$*]: No sistema indicado na Fig. 9.36, se a função de transferência de malha aberta $G(s)H(s)$ possuir k pólos no semiplano direito do plano s e $\lim_{s \to \infty} G(s)H(s) =$ constante, então, para a estabilidade, o lugar geométrico de $G(j\omega)H(j\omega)$, conforme ω varia de $-\infty$ até ∞, deve envolver o ponto $-1 + j0$ k vezes no sentido anti-horário.

Considerações sobre o critério de estabilidade de Nyquist

1. Este critério pode ser expresso como

$$Z = N + P$$

onde

Z = número de zeros de $1 + G(s)H(s)$ no semiplano direito do plano s
N = número de envolvimentos do ponto $-1 + j0$ no sentido horário
P = número de pólos de $G(s)H(s)$ no semiplano direito do plano s

Se P for diferente de zero, para um sistema de controle estável devemos ter $Z = 0$, ou $N = -P$, o que significa que devemos ter P envolvimentos do ponto $-1 + j0$ no sentido anti-horário.

Se $G(s)H(s)$ não possuir qualquer pólo no semiplano direito do plano s, então $Z = N$. Portanto, para estabilidade não devemos ter nenhum envolvimento do ponto $-1 + j0$ pelo lugar geométrico de $G(j\omega)H(j\omega)$. Neste caso, não é necessário considerar o lugar geométrico para todo o eixo $j\omega$, porém, somente a parte de freqüência positiva. A estabilidade de um sistema deste tipo pode ser determinada verificando-se se o ponto $-1 + j0$ é circundado pelo gráfico de Nyquist de $G(j\omega)H(j\omega)$. A região envolvida pelo gráfico de Nyquist é indicada na Fig. 9.40. Para estabilidade, o ponto $-1 + j0$ deve estar do lado de fora da região hachurada.

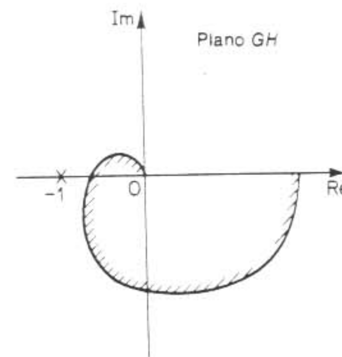


Fig. 9.40 Região envolvida por um gráfico de Nyquist.

2. Devemos tomar cuidado quando verificarmos a estabilidade de sistemas de múltiplas malhas desde que estas podem incluir pólos no semiplano direito do plano s. (Note que embora uma malha-interna possa ser instável, o sistema de malha-fechada como um todo pode ser estável, através de um projeto apropriado.) A simples inspeção dos envolvimentos do ponto $-1 + j0$ pelo lugar geométrico de $G(j\omega)H(j\omega)$ não é suficiente para detectar instabilidade em sistemas de múltiplas malhas. Nestes casos, entretanto, a existência ou não de qualquer pólo de $1 + G(s)H(s)$ no semiplano direito do plano s pode ser determinada facilmente aplicando-se o critério de estabilidade de Routh para o denominador de $G(s)H(s)$.

Se funções transcendentais, tais como atraso de transporte e^{-Ts}, forem incluídas em $G(s)H(s)$, elas devem ser aproximadas, por meio de uma expansão em série, antes de ser aplicado o critério de estabilidade de Routh. Uma forma da expansão em série de e^{-Ts} é dada a seguir:

$$e^{-Ts} = \frac{1 - \dfrac{Ts}{2} + \dfrac{(Ts)^2}{8} - \dfrac{(Ts)^3}{48} + \cdots}{1 + \dfrac{Ts}{2} + \dfrac{(Ts)^2}{8} + \dfrac{(Ts)^3}{48} + \cdots}$$

Em primeira aproximação, podemos considerar apenas os dois primeiros termos no numerador e denominador, respectivamente, ou

$$e^{-Ts} \doteq \frac{1 - \dfrac{Ts}{2}}{1 + \dfrac{Ts}{2}} = \frac{2 - Ts}{2 + Ts}$$

Isto dá uma boa aproximação para o atraso de transporte em uma faixa de freqüên-
cia de $0 \leq \omega \leq (0.5/T)$. [Note que o módulo de $(2 - j\omega T)/(2 + j\omega T)$ é sempre igual a
um, e o ângulo de fase de $(2 - j\omega T)/(2 + j\omega T)$ é aproximadamente igual àquele do
atraso de transporte, dentro da faixa de freqüência estabelecida.]

3. Se o lugar geométrico de $G(j\omega)H(j\omega)$ passa pelo ponto $-1 + j0$, então os
zeros da equação característica, ou os pólos de malha-fechada, estão localizados
sobre o eixo $j\omega$. Isto não é desejável para sistemas de controle práticos. Para um
sistema de malha-fechada bem projetado, nenhuma raiz da equação característica
deve estar sobre o eixo $j\omega$.

Caso especial em que $G(s)H(s)$ contém pólos e/ou zeros sobre o eixo $j\omega$. Na
discussão anterior, supusemos que a função de transferência de malha-aberta
$G(s)H(s)$ não possuía pólos ou zeros na origem. Consideraremos agora o caso em
que $G(s)H(s)$ contém pólos e/ou zeros sobre o eixo $j\omega$.

Desde que o caminho de Nyquist não passa por pólos ou zeros de $G(s)H(s)$, se
a função $G(s)H(s)$ possui pólos ou zeros na origem (ou sobre o eixo $j\omega$ em pontos
diferentes da origem), o contorno no plano s deve ser modificado. O modo usual de
modificação do contorno próximo à origem é usar uma semicircunferência com um
raio infinitesimal ϵ, conforme indicado na Fig. 9.41. Um ponto representativo s
move-se ao longo do eixo negativo de $j\omega$ desde $-j\infty$ até $-j0_-$. De $s = j0_-$ até $s =$
$j0_+$, o ponto se move ao longo da semicircunferência de raio ϵ (onde $\epsilon \ll 1$), e então
move-se ao longo do eixo $j\omega$ positivo desde $j0_+$ até $j\infty$. De $s = j\infty$, o contorno segue
uma semicircunferência com raio infinito e o ponto representativo retorna ao ponto
inicial. A área que o contorno fechado modificado evita é muito pequena e tende a
zero, conforme o raio ϵ tende a zero. Portanto, todos os pólos e zeros, se houver, no
semiplano direito do plano s serão envolvidos por este contorno.

Considere, por exemplo, um sistema de malha-fechada cuja função de transfe-
rência de malha-aberta é dada por

$$G(s)H(s) = \frac{K}{s(Ts + 1)}$$

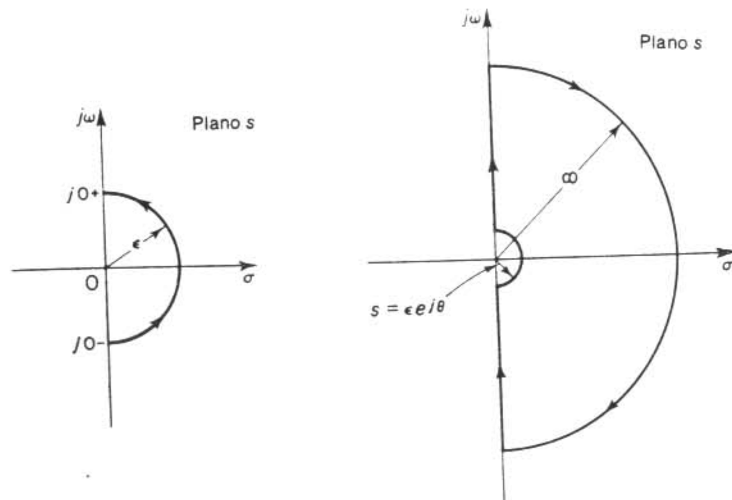


Fig. 9.41 Contornos fechados no plano s evitando pólos e zeros na origem.

Os pontos correspondentes a $s = j0_+$ e $s = j0_-$ sobre o lugar geométrico de $G(s)H(s)$
no plano $G(s)H(s)$ são $-j\infty$ e $j\infty$, respectivamente. Sobre o caminho semicircular
com raio ϵ (onde $\epsilon \ll 1$), a variável complexa s pode ser escrita

$$s = \epsilon e^{j\theta}$$

onde θ varia de $-90°$ para $+90°$. Então, $G(s)H(s)$ resulta em

$$G(\epsilon e^{j\theta})H(\epsilon e^{j\theta}) = \frac{K}{\epsilon e^{j\theta}} = \frac{K}{\epsilon} e^{-j\theta}$$

O valor de K/ϵ tende a infinito conforme ϵ tende a zero, e $-\theta$ varia de $90°$ a $-90°$
conforme um ponto representativo s move-se ao longo da semicircunferência.
Portanto, os pontos $G(j0_-)H(j0_-) = j\infty$ e $G(j0_+)H(j0_+) = -j\infty$ são unidos por uma
semicircunferência de raio infinito no semiplano direito do plano GH. A semicir-
cunferência infinitesimal em torno da origem mapeia no plano GH em uma semicir-
cunferência de raio infinito. A Fig. 9.42 mostra o contorno no plano s e o lugar
geométrico de $G(s)H(s)$ no plano GH. Os pontos A, B e C do contorno no plano s
mapeiam nos pontos respectivos A', B' e C' do lugar geométrico $G(s)H(s)$. Como
pode ser observado da Fig. 9.42, os pontos D, E e F na semicircunferência de raio
infinito no plano s mapeiam na origem do plano GH. Desde que não há pólo no
semiplano direito do plano s e o lugar geométrico de $G(s)H(s)$ não envolve o ponto
$-1 + j0$, não há zeros da função $1 + G(s)H(s)$ no semiplano direito do plano s.
Portanto, o sistema é estável.

Para uma função de transferência de malha-aberta $G(s)H(s)$ contendo um fator
$1/s^n$ (onde $n = 2, 3, \ldots$), o gráfico de $G(s)H(s)$ possui n semicircunferências de raio
infinito no sentido horário em torno da origem, conforme um ponto representativo s

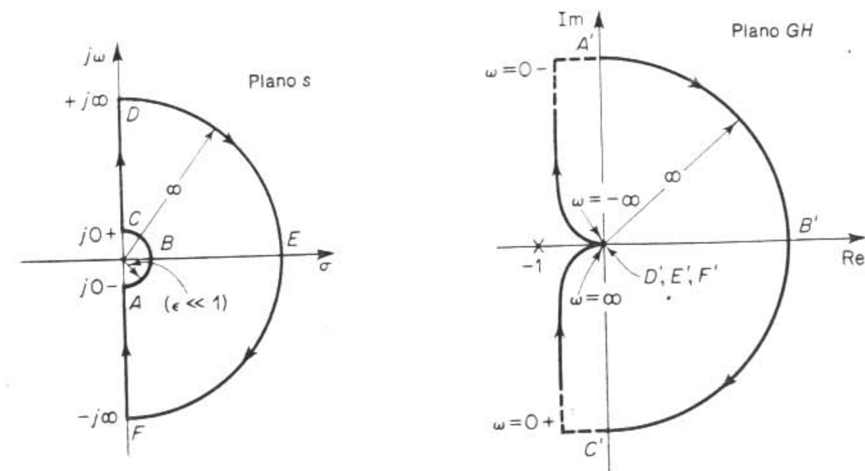


Fig. 9.42 Contorno no plano s e o lugar geométrico de $G(s)H(s)$ no plano GH onde $G(s)H(s) =$
$K/[s(Ts + 1)]$.

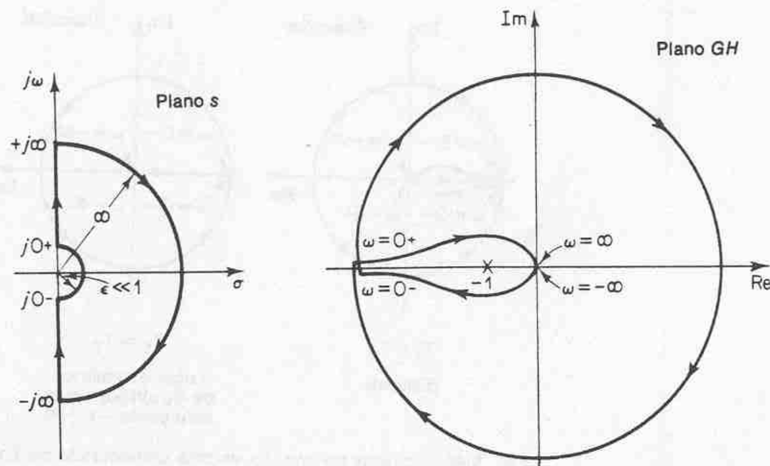


Fig. 9.43 Contorno no plano s e o lugar geométrico de $G(s)H(s)$ no plano GH onde $G(s)H(s) = K/[s^2(Ts+1)]$.

percorre a semicircunferência de raio ϵ (onde $\epsilon \ll 1$). Por exemplo, considere a seguinte função de transferência de malha-aberta:

$$G(s)H(s) = \frac{K}{s^2(Ts+1)}$$

Então

$$\lim_{s \to \epsilon e^{j\theta}} G(s)H(s) = \frac{K}{\epsilon^2 e^{2j\theta}} = \frac{K}{\epsilon^2} e^{-2j\theta}$$

Conforme θ varia desde $-90°$ a $90°$ no plano s, o ângulo de $G(s)H(s)$ varia desde $180°$ até $-180°$, conforme a Fig. 9.43. Desde que não há pólos no semiplano direito do plano s e o lugar geométrico envolve o ponto $-1+j0$ duas vezes no sentido horário, para qualquer valor positivo de K, há dois zeros de $1+G(s)H(s)$ no semiplano direito do plano s. Portanto, este sistema é sempre instável.

Note que uma análise equivalente pode ser realizada se $G(s)H(s)$ contiver pólos e/ou zeros sobre o eixo $j\omega$. O critério de estabilidade de Nyquist pode agora ser generalizado como segue:

Critério de estabilidade de Nyquist (Para o caso geral onde $G(s)H(s)$ contém pólos e/ou zeros no eixo $j\omega$.): No sistema indicado na Fig. 9.36, se a função de transferência de malha-aberta $G(s)H(s)$ possuir k pólos no semiplano direito do plano s, então, para estabilidade, o lugar geométrico de $G(s)H(s)$ deve envolver o ponto $-1+j0$ k vezes no sentido anti-horário, conforme um ponto representativo s percorre o caminho modificado de Nyquist no sentido horário.

9.6 ANÁLISE DE ESTABILIDADE

Nesta seção, apresentaremos alguns exemplos ilustrativos da análise de estabilidade de sistemas de controle utilizando o critério de estabilidade de Nyquist.

Se o caminho de Nyquist no plano s envolver Z zeros e P pólos de $1+G(s)H(s)$ e não passar por nenhum pólo ou zero de $1+G(s)H(s)$, conforme um ponto representativo s percorre o caminho de Nyquist no sentido horário, então o contorno correspondente no plano $G(s)H(s)$ envolve o ponto $-1+j0$ $N = Z - P$ vezes no sentido horário. (Valores negativos de N implicam envolvimentos no sentido anti-horário.)

No exame da estabilidade de sistemas de controle lineares utilizando o critério de estabilidade de Nyquist, verificamos que podem ocorrer três possibilidades:

1. Não há envolvimento do ponto $-1+j0$. Isto implica que o sistema é estável se não houver pólos de $G(s)H(s)$ no semiplano direito do plano s; caso contrário, o sistema é instável.
2. Há um envolvimento ou envolvimentos do ponto $-1+j0$ no sentido anti-horário. Neste caso o sistema é estável se o número de envolvimentos no sentido anti-horário for o mesmo que o número de pólos de $G(s)H(s)$ no semiplano direito do plano s; caso contrário o sistema é instável.
3. Há um envolvimento ou envolvimentos no sentido horário do ponto $-1+j0$. Neste caso o sistema é instável.

Nos exemplos a seguir, suporemos que os valores do ganho K e as constantes de tempo (tais como T, T_1 e T_2) sejam todos positivos.

Exemplo 9.6 Considere um sistema de malha-fechada cuja função de transferência de malha-aberta é dada por

$$G(s)H(s) = \frac{K}{(T_1 s+1)(T_2 s+1)}$$

Examine a estabilidade do sistema.

Um gráfico de $G(j\omega)H(j\omega)$ é indicado na Fig. 9.44. Desde que $G(s)H(s)$ não possui

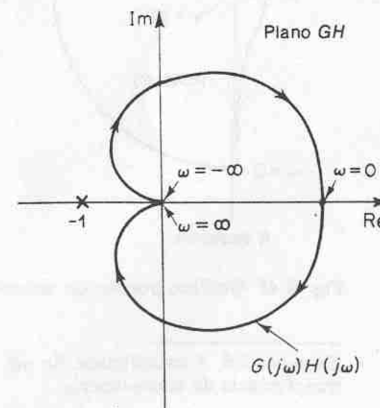


Fig. 9.44 Gráfico polar de $G(j\omega)H(j\omega)$ considerada no Exemplo 9.6.

qualquer pólo no semiplano direito do plano s e o ponto $-1 + j0$ não é envolvido pelo lugar geométrico de $G(j\omega)H(j\omega)$, este sistema é estável para quaisquer valores positivos de K, T_1 e T_2.

Exemplo 9.7 Considere o sistema com a seguinte função de transferência de malha-aberta:

$$G(s)H(s) = \frac{K}{s(T_1s + 1)(T_2s + 1)}$$

Determine a estabilidade do sistema para os dois casos, (1) o ganho K é pequeno, (2) K é grande.

Os gráficos de Nyquist da função de transferência de malha-aberta com um pequeno valor de K e com um grande valor de K são indicados na Fig. 9.45. O número de pólos de $G(s)H(s)$ no semiplano direito do plano s é zero. Portanto, para este sistema ser estável, é necessário que $N = Z = 0$ ou que o lugar geométrico de $G(s)H(s)$ não envolva o ponto $-1 + j0$.

Para pequenos valores de K, não há envolvimento do ponto $-1 + j0$. Conseqüentemente, o sistema é estável para pequenos valores de K. Para grandes valores de K, o lugar geométrico de $G(s)H(s)$ envolve o ponto $-1 + j0$ duas vezes no sentido horário, indicando dois pólos de malha-fechada no semiplano direito do plano s e a instabilidade do sistema. (Para boa precisão, K deve ser grande. Do ponto de vista de estabilidade, entretanto, um valor grande de K causa uma estabilidade pobre ou mesmo instabilidade. A fim de estabelecer um compromisso entre precisão e estabilidade é necessário inserir uma rede de compensação no sistema. As técnicas de compensação no domínio de freqüência são discutidas no Cap. 10.)

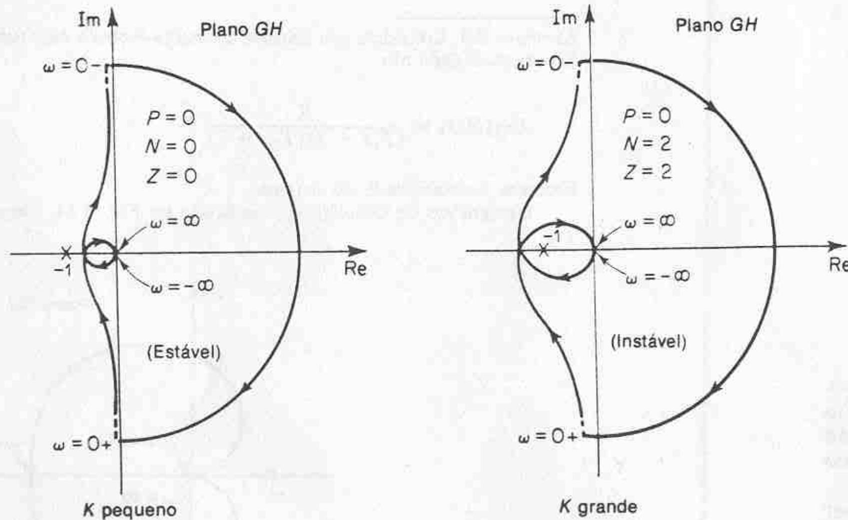


Fig. 9.45 Gráficos polares do sistema considerado no Exemplo 9.7.

Exemplo 9.8 A estabilidade de um sistema de malha-fechada com a seguinte função de transferência de malha-aberta

$$G(s)H(s) = \frac{K(T_2s + 1)}{s^2(T_1s + 1)}$$

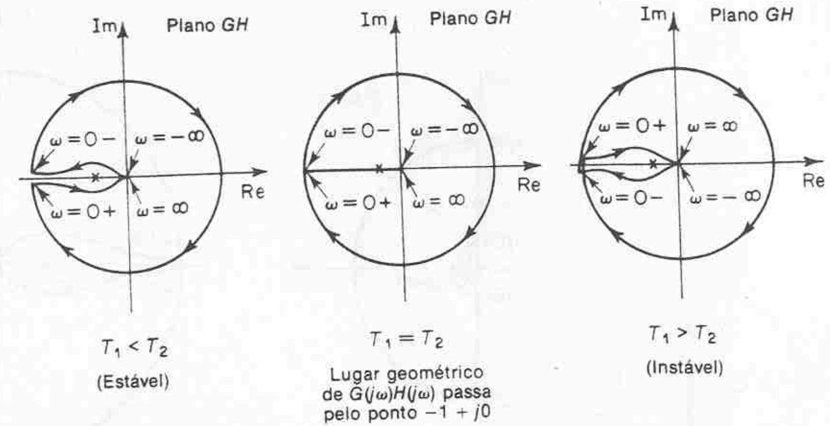


Fig. 9.46 Gráficos polares do sistema considerado no Exemplo 9.8.

depende dos valores relativos de T_1 e T_2. Desenhe os gráficos de Nyquist e determine a estabilidade do sistema.

Os gráficos dos lugares geométricos de $G(s)H(s)$ para os três casos, $T_1 < T_2$, $T_1 = T_2$ e $T_1 > T_2$, são indicados na Fig. 9.46. Para $T_1 < T_2$, o lugar geométrico de $G(s)H(s)$ não envolve o ponto $-1 + j0$, e o sistema de malha-fechada é estável. Para $T_1 = T_2$, o lugar geométrico de $G(s)H(s)$ passa pelo ponto $-1 + j0$, o que indica que os pólos de malha-fechada estão localizados sobre o eixo $j\omega$. Para $T_1 > T_2$, o lugar geométrico de $G(s)H(s)$ envolve o ponto $-1 + j0$ duas vezes no sentido horário. Portanto, o sistema de malha-fechada possui dois pólos de malha-fechada no semiplano direito do plano s, e o sistema é instável.

Exemplo 9.9 Considere o sistema de malha-fechada com a seguinte função de transferência de malha-aberta:

$$G(s)H(s) = \frac{K}{s(Ts - 1)}$$

Determine a estabilidade do sistema.

A função $G(s)H(s)$ tem um pólo ($s = 1/T$) no semiplano direito do plano s. Portanto, $P = 1$. O gráfico de Nyquist mostrado na Fig. 9.47 indica que o lugar geométrico de $G(s)H(s)$ envolve o ponto $-1 + j0$ uma vez no sentido horário. Portanto, $N = 1$. Desde que $Z = N + P$, verificamos que $Z = 2$. Isto significa que o sistema de malha-fechada possui dois pólos de malha-fechada no semiplano direito do plano s e é instável.

Exemplo 9.10 **Investigue a estabilidade de um sistema de malha-fechada com a seguinte função de transferência de malha-aberta:**

$$G(s)H(s) = \frac{K(s + 3)}{s(s - 1)}$$

A função de transferência de malha-aberta possui um pólo ($s = 1$) no semiplano direito do plano s, ou $P = 1$. O sistema de malha-aberta é instável. O gráfico de Nyquist mostrado na Fig. 9.48 indica que o ponto $-1 + j0$ é envolvido pelo lugar geométrico de $G(s)H(s)$ uma vez no sentido anti-horário. Portanto, $N = 1$. Conseqüentemente Z, determinado a partir de $Z = N +$

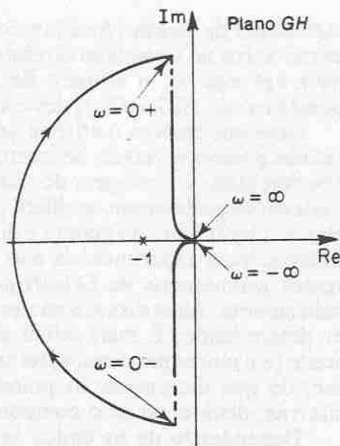


Fig. 9.47 Gráfico polar do sistema considerado no Exemplo 9.9.

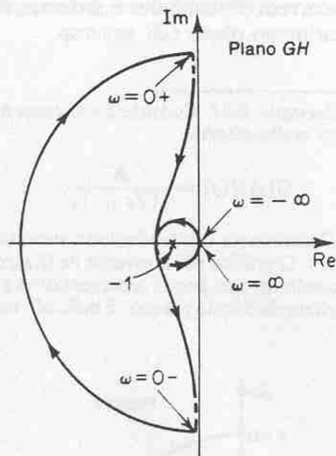


Fig. 9.48 Gráfico polar do sistema considerado no Exemplo 9.10.

P, é igual a zero, o que indica que não há nenhum zero de $1 + G(s)H(s)$ no semiplano direito do plano s, e o sistema de malha-fechada é estável. Este é um dos exemplos onde um sistema de malha-aberta instável torna-se estável quando a malha é fechada.

Sistemas condicionalmente estáveis. A Fig. 9.49 mostra um exemplo de um lugar geométrico de $G(j\omega)H(j\omega)$ para o qual o sistema de malha-fechada pode ser feito instável variando-se o ganho de malha-aberta. Se o ganho de malha-aberta for aumentado suficientemente, o lugar geométrico de $G(j\omega)H(j\omega)$ envolve o ponto $-1 + j0$ duas vezes, e o sistema resulta instável. Se o ganho de malha-aberta é

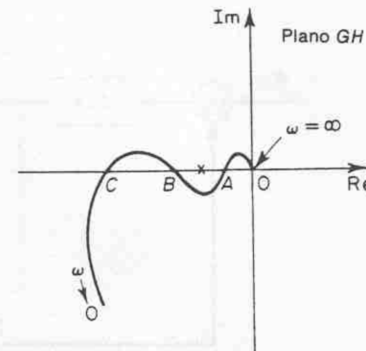


Fig. 9.49 Gráfico polar de um sistema condicionalmente estável.

diminuído suficientemente, ainda o lugar geométrico de $G(j\omega)H(j\omega)$ envolve o ponto $-1 + j0$ duas vezes. O sistema é estável apenas para a faixa limitada de valores do ganho de malha-aberta para os quais o ponto $-1 + j0$ está completamente do lado de fora do lugar geométrico de $G(j\omega)H(j\omega)$. Um sistema deste tipo é dito condicionalmente estável.

Um sistema condicionalmente estável é estável para os valores do ganho de malha-aberta que estão entre os valores críticos, mas é instável se o ganho de malha-aberta for aumentado ou diminuído suficientemente. Um sistema deste tipo resulta instável quando são aplicados grandes sinais na entrada já que um sinal grande pode causar saturação, o que por sua vez reduz o ganho de malha-aberta do sistema. É aconselhável evitar estas situações visto que o sistema pode tornar-se instável, se o ganho de malha-aberta cair abaixo de um valor crítico.

Para operação estável do sistema condicionalmente estável aqui considerado, o ponto crítico $-1 + j0$ deve estar localizado nas regiões entre OA e BC indicadas na Fig. 9.49.

Sistemas de múltiplas malhas. Considere o sistema indicado na Fig. 9.50. Este é um sistema de múltiplas malhas. A malha interna possui a função de transferência

$$G(s) = \frac{G_2(s)}{1 + G_2(s)H_2(s)}$$

Se $G(s)$ é instável, os efeitos da instabilidade são para produzir um pólo ou pólos no semiplano direito do plano s. Então a equação característica da malha interna, $1 + G_2(s)H_2(s) = 0$, possui um zero ou zeros nesta parte do plano. Se $G_2(s)$ e $H_2(s)$ possuírem nesta parte P_1 pólos, então o número Z_1 de zeros no semiplano direito de $1 + G_2(s)H_2(s)$ pode ser determinado de $Z_1 = N_1 + P_1$, onde N_1 é o número de envolvimentos do ponto $-1 + j0$, no sentido horário, pelo lugar geométrico de $G_2(s)H_2(s)$. Desde que a função de transferência de malha-aberta do sistema inteiro é dada por $G_1(s)G(s)H_1(s)$, a estabilidade deste sistema de malha-fechada pode ser determinada do gráfico de Nyquist de $G_1(s)G(s)H_1(s)$ e o conhecimento dos pólos no semiplano direito de $G_1(s)G(s)H_1(s)$.

Note que se for eliminada uma malha de realimentação por meio de reduções de diagrama de blocos, há uma possibilidade de serem introduzidos pólos instáveis;

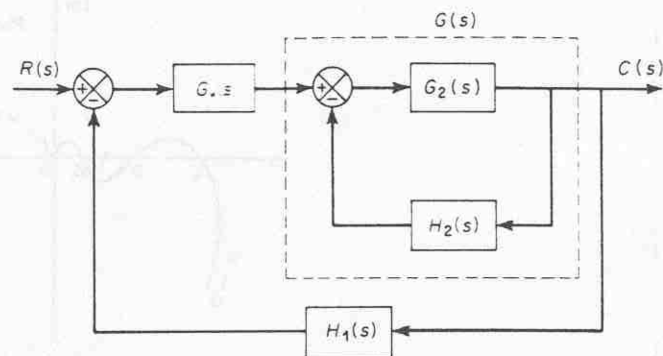


Fig. 9.50 Sistema de múltiplas malhas.

se o ramo direito for eliminado por meio de reduções de diagrama de blocos, há uma possibilidade de serem introduzidos zeros no semiplano direito. Portanto, devemos notar todos os pólos e zeros do semiplano direito conforme eles aparecem de reduções de malhas intermediárias. Este conhecimento é necessário na determinação da estabilidade de sistemas de múltiplas malhas.

Na análise de sistemas de múltiplas malhas deste tipo, a função de transferência inversa pode algumas vezes ser utilizada de modo a permitir análise gráfica; isto evita muito cálculo numérico.

Critério de estabilidade de Nyquist aplicado a gráficos polares inversos. Na análise anterior, o critério de estabilidade de Nyquist foi aplicado aos gráficos polares da função de transferência de malha-aberta $G(s)H(s)$. O critério de estabilidade de Nyquist pode igualmente ser aplicado aos gráficos polares inversos. A dedução matemática do critério de estabilidade de Nyquist para gráficos polares inversos é idêntica à efetuada para gráficos polares diretos.

O gráfico polar inverso de $G(j\omega)H(j\omega)$ é um gráfico de $1/[G(j\omega)H(j\omega)]$ em função de ω. Por exemplo, se $G(j\omega)H(j\omega)$ é

$$G(j\omega)H(j\omega) = \frac{j\omega T}{1 + j\omega T}$$

então

$$\frac{1}{G(j\omega)H(j\omega)} = \frac{1}{j\omega T} - 1$$

O gráfico polar inverso para $\omega \geq 0$ é a metade inferior da reta vertical iniciando no ponto (1; 0) sobre o eixo real.

O critério de estabilidade de Nyquist aplicado aos gráficos polares inversos pode ser estabelecido como segue: Para um sistema de malha-fechada ser estável, o envolvimento, se houver, do ponto $-1 + j0$ pelo lugar geométrico de $1/[G(s)H(s)]$ (conforme s percorre o caminho de Nyquist) deve ser anti-horário, e o número destes envolvimentos deve ser igual ao número de pólos de $1/[G(s)H(s)]$ [isto é, os zeros de $G(s)H(s)$] que estão no semiplano direito do plano s. (O número de zeros de $G(s)H(s)$ no semiplano direito do plano s pode ser determinado pelo critério de

estabilidade de Routh.) Se a função de transferência de malha-aberta $G(s)H(s)$ não possuir zeros no semiplano direito do plano s, então, para um sistema de malha-fechada ser estável, o número de envolvimentos do ponto $-1 + j0$ pelo lugar geométrico de $1/[G(s)H(s)]$ deve ser zero.

Note que embora o critério de estabilidade de Nyquist possa ser aplicado aos gráficos polares inversos, se forem incorporados dados de resposta em freqüência experimentais, a contagem do número de envolvimentos do lugar geométrico de $1/[G(s)H(s)]$ pode tornar-se difícil porque a defasagem correspondente ao caminho semicircular infinito no plano s é difícil de medir. Em geral, se os dados de resposta em freqüência experimentais não puderem ser colocados em forma analítica, os lugares geométricos de $G(j\omega)H(j\omega)$ e de $1/[G(j\omega)H(j\omega)]$ devem ser construídos graficamente. Além disso, o número de zeros de $G(s)H(s)$ no semiplano direito deve ser determinado. É mais difícil determinar os zeros de $G(s)H(s)$ no semiplano direito (em outras palavras, determinar se um dado componente é ou não de mínima fase) do que determinar os pólos de $G(s)H(s)$ no semiplano direito (em outras palavras, determinar se o componente é ou não estável).

Dependendo de os dados serem analíticos ou gráficos e dos componentes serem ou não de mínima fase, um teste de estabilidade apropriado deve ser utilizado para sistemas de múltiplas malhas. Se os dados forem fornecidos em forma analítica, ou se as expressões matemáticas de todos os componentes forem conhecidas, a aplicação do critério de estabilidade de Nyquist para gráficos polares inversos não acarreta dificuldades e sistemas de múltiplas malhas podem ser analisados e projetados no plano GH inverso.

Exemplo 9.11 Considere o sistema de malha-fechada com a seguinte função de transferência de malha-aberta:

$$G(s)H(s) = \frac{K}{s(Ts + 1)}$$

Determine a estabilidade do sistema utilizando o gráfico polar inverso.

O gráfico polar inverso de $G(s)H(s)$, isto é, o gráfico do lugar geométrico de $1/[G(s)H(s)]$ conforme um ponto representativo s percorre o contorno fechado envolvendo todo o semiplano direito do plano s é indicado na Fig. 9.51, junto com o contorno no plano s. Conforme s

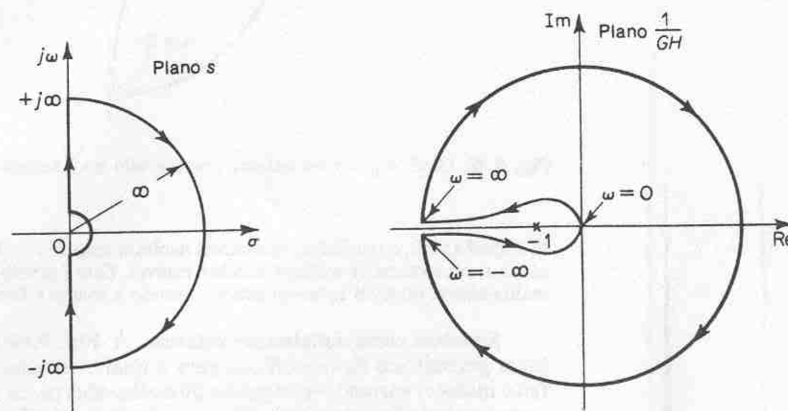


Fig. 9.51 Contorno no plano s e lugar geométrico de $G(s)H(s)$ no plano $1/GH$ onde $G(s)H(s) = K/[s(Ts + 1)]$.

percorre o contorno fechado no sentido horário, não é envolvido o ponto $-1 + j0$. Portanto, o sistema de malha-fechada é estável. Neste sistema, a origem do plano s é mapeada na origem do plano $1/GH$. A semicircunferência com raio infinito, ou $s = \infty e^{j\theta}$, onde $-90° \leq \theta \leq 90°$, corresponde a uma circunferência de raio infinito no plano $1/GH$.

Note que, se a função de transferência de malha-fechada $G(s)H(s)$ possui atraso de transporte, por exemplo,

$$G(s)H(s) = \frac{Ke^{-j\omega L}}{s(Ts+1)}$$

então o número de envolvimentos do ponto $-1 + j0$ pelo lugar geométrico de $1/[G(s)H(s)]$ é infinito, e o critério de estabilidade de Nyquist não pode ser aplicado para o gráfico polar inverso desta função de transferência de malha-aberta.

Análise de estabilidade relativa através dos caminhos de Nyquist modificados. O caminho de Nyquist para testes de estabilidade pode ser modificado de modo que permita a investigação da estabilidade relativa de sistemas de malha-fechada. Para a seguinte equação característica de segunda-ordem,

$$s^2 + 2\zeta\omega_n s + \omega_n^2 = 0 \qquad (0 < \zeta < 1)$$

as raízes são complexas conjugadas com valores

$$s_1 = -\zeta\omega_n + j\omega_n\sqrt{1-\zeta^2}$$
$$s_2 = -\zeta\omega_n - j\omega_n\sqrt{1-\zeta^2}$$

Se estas raízes forem colocadas em um gráfico no plano s, como indicado na Fig. 9.52, então verificamos que sen $\theta = \zeta$, ou seja, o ângulo θ é indicativo da relação de amortecimento ζ. Conforme θ torna-se menor, também o será o valor de ζ.

Se modificarmos o caminho de Nyquist e utilizarmos retas radiais com ângulo θ_x ao invés do eixo $j\omega$, conforme indicado na Fig. 9.53, então pode ser dito, seguindo os mesmos motivos do caso do critério de estabilidade de Nyquist, que se o lugar geométrico de $G(s)H(s)$ correspondente ao contorno no plano s modificado não envolver o ponto $-1 + j0$ e nenhum dos pólos de $G(s)H(s)$ estiver dentro do contorno do plano s fechado, então este contorno não envolve qualquer zero de $1 + G(s)H(s)$. A equação característica, $1 + G(s)H(s) = 0$, então, não possui qualquer raiz dentro do contorno no plano s modificado. Se nenhum pólo de malha-fechada

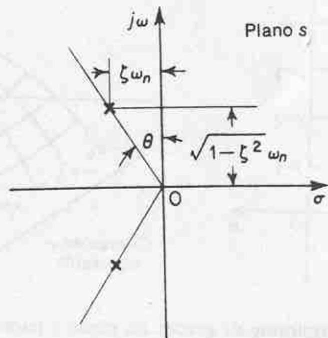


Fig. 9.52 Gráfico das raízes complexas conjugadas no plano s.

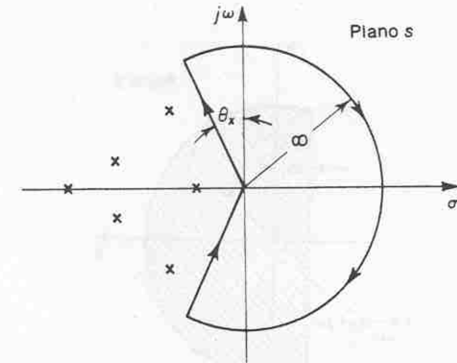


Fig. 9.53 Caminho de Nyquist modificado.

de um sistema de maior ordem for envolvido por este contorno, podemos afirmar que a relação de amortecimento de cada par de pólos complexos conjugados de malha-fechada do sistema é maior do que sen θ_x.

O lugar geométrico de $G(s)H(s)$ pode ser construído substituindo-se s por

$$-\zeta_x\omega + j\omega\sqrt{1-\zeta_x^2}$$

onde ζ_x = sen θ_x, com ω tomando todos os valores desde $-\infty$ até ∞, e esboçando o lugar geométrico resultante. O lugar geométrico para

$$G(-\zeta_x\omega - j\omega\sqrt{1-\zeta_x^2})H(-\zeta_x\omega - j\omega\sqrt{1-\zeta_x^2})$$

é a imagem especular em relação ao eixo real do lugar geométrico para

$$G(-\zeta_x\omega + j\omega\sqrt{1-\zeta_x^2})H(-\zeta_x\omega + j\omega\sqrt{1-\zeta_x^2})$$

Um diagrama deste tipo é denominado diagrama de Vazsonyi. O diagrama de Nyquist é um caso especial deste último.

Analogamente, se o contorno no plano s consiste em uma reta à esquerda e paralela ao eixo $j\omega$ e a uma distância $-\sigma_0$ (ou a reta $s = -\sigma_0 + j\omega$) e a semicircunferência de raio infinito envolvendo todo o semiplano direito do plano s e aquela parte do semiplano esquerdo do plano s entre as retas $s = -\sigma_0 + j\omega$ e $s = j\omega$, conforme indicado na Fig. 9.54(a), então podemos estabelecer o seguinte: Se o lugar geométrico de $G(s)H(s)$ correspondente a este contorno no plano s não envolver o ponto $-1 + j0$ e $G(s)H(s)$ não possuir pólos dentro do contorno fechado no plano s, então a equação característica não possuirá qualquer zero na região envolvida pelo contorno do plano s modificado. Todas as raízes da equação característica estarão à esquerda da reta $s = -\sigma_0 + j\omega$. Um exemplo de um lugar geométrico de $G(-\sigma_0 + j\omega)H(-\sigma_0 + j\omega)$ conjuntamente com um lugar geométrico de $G(j\omega)H(j\omega)$ é indicado na Fig. 9.54(b). O valor $1/\sigma_0$ é indicativo da constante de tempo dos pólos de malha-fechada dominantes. Se todas as raízes estiverem do lado de fora do contorno no plano s, todas as constantes de tempo da função de transferência de malha-fechada serão menores do que $1/\sigma_0$. Se o contorno no plano s é escolhido conforme indicado na Fig. 9.55, então o teste dos envolvimentos do ponto $-1 + j0$

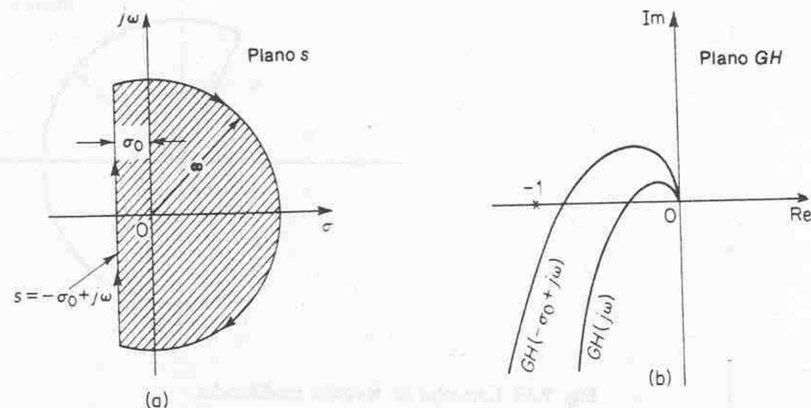


Fig. 9.54(a) Caminho de Nyquist modificado; (b) gráficos polares do lugar geométrico de $G(-\sigma_0 + j\omega)H(-\sigma_0 + j\omega)$ e do lugar geométrico de $G(j\omega)H(j\omega)$ no plano GH.

revela a existência ou não das raízes da equação característica do sistema de malha-fechada dentro do contorno no plano s. Se o teste revela que nenhuma raiz está dentro do contorno no plano s, então é claro que todos os pólos de malha-fechada possuirão relações de amortecimento maiores do que ζ_r e constantes de tempo menores do que $1/\sigma_0$. Portanto, considerando um contorno apropriado no plano s, podemos investigar constantes de tempo e relações de amortecimento de pólos de malha-fechada de funções de transferência de malha-aberta conhecidas.

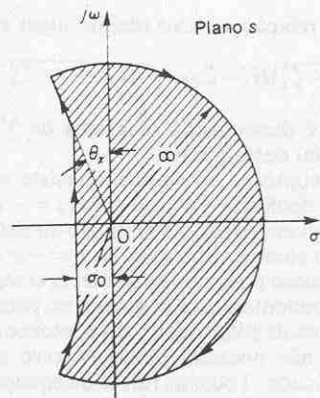


Fig. 9.55 Caminho de Nyquist modificado.

9.7 ESTABILIDADE RELATIVA

No projeto de sistema de controle, exige-se que o sistema seja estável. Além disso, é necessário que o sistema possua uma adequada estabilidade relativa.

Nesta seção, mostraremos como o gráfico de Nyquist indica não apenas se o sistema é ou não estável mas também o grau de estabilidade de um sistema estável.

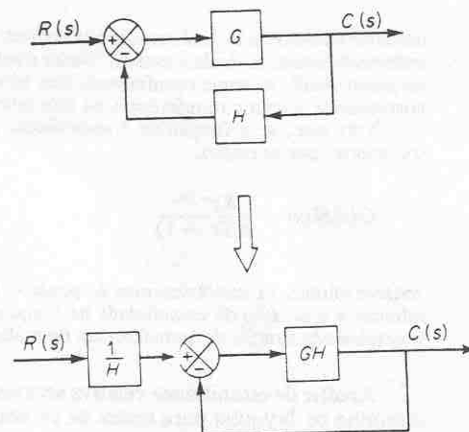


Fig. 9.56 Modificação de um sistema com elementos de realimentação para um sistema com realimentação unitária.

O gráfico de Nyquist também fornece informação de como a estabilidade pode ser melhorada, se isto for necessário. (Para detalhes, vide Cap. 10).

Na discussão a seguir, suporemos que os sistemas considerados possuem realimentação unitária. Note que é sempre possível reduzir um sistema com elementos de realimentação para um sistema com realimentação unitária, conforme indicado na Fig. 9.56. Conseqüentemente é possível estender-se a análise de estabilidade relativa para sistema com realimentação unitária para um outro sistema com realimentação não unitária.

Consideraremos também que, salvo menção em contrário, os sistemas sejam de mínima fase; isto é, a função de transferência de malha-aberta $G(s)$ não possui pólos ou zeros no semiplano direito do plano s.

Análise de estabilidade relativa via mapeamento conforme. Um dos importantes problemas na análise de um sistema de controle é determinar todos os pólos de

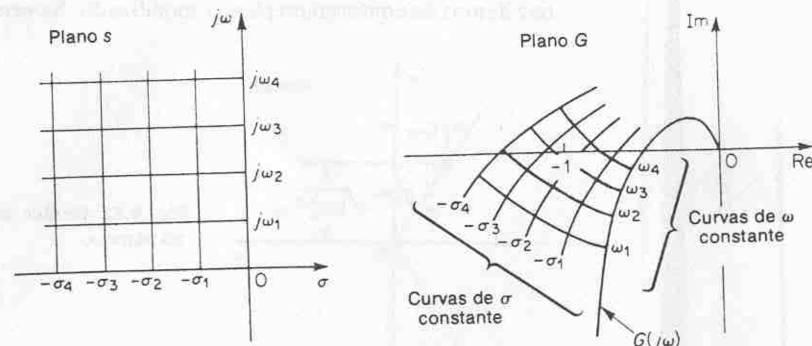


Fig. 9.57 Mapeamento conforme de grades no plano s para o plano $G(s)$.

malha-fechada ou pelo menos aqueles mais próximos do eixo $j\omega$ (ou os pólos de malha-fechada correspondentes ao par dominante.) Se as características de resposta em freqüência de malha-aberta de um sistema forem conhecidas, pode ser possível estimar os pólos de malha-fechada mais próximos do eixo $j\omega$. A técnica a ser aqui apresentada é essencialmente gráfica e baseada em um mapeamento conforme do plano s no plano $G(s)$.

Considere o mapeamento conforme das retas de σ constante (retas $s = \sigma + j\omega$, onde σ é constante e ω varia) e retas de ω constante (retas $s = \sigma + j\omega$, onde ω é constante e σ varia) no plano s para o plano $G(s)$. A reta $\sigma = 0$ (isto é, o eixo $j\omega$) no plano s mapeia o gráfico de Nyquist no plano $G(s)$. As retas de σ constante no plano s mapeiam em curvas que são similares ao gráfico de Nyquist e, sob certo sentido, paralelas ao gráfico de Nyquist como indicado na Fig. 9.57. As retas de ω constante no plano s mapeiam em curvas, também indicadas na Fig. 9.57.

Embora as formas dos lugares geométricos de σ constante e ω constante no plano $G(s)$ e a proximidade da abordagem do lugar geométrico de $G(j\omega)$ para o ponto $-1 + j0$ dependam de um $G(s)$ particular, a proximidade da abordagem do lugar geométrico de $G(j\omega)$ para o ponto $-1 - j0$ é uma indicação da estabilidade relativa de um sistema estável. Em geral, podemos esperar que quanto mais próximo o lugar geométrico de $G(j\omega)$ estiver em relação ao ponto $-1 + j0$, maior será a sobrelevação máxima na resposta transitória ao degrau e mais longo será o tempo para o amortecimento.

Considere os dois sistemas indicados nas Figs. 9.58(a) e (b). (Na Fig. 9.58 as cruzes indicam pólos de malha-fechada.) O sistema (a) é obviamente mais estável do que o sistema (b) porque os pólos de malha-fechada do sistema (a) estão localizados mais à esquerda daqueles correspondentes ao sistema (b). As Figs. 9.59(a) e (b) mostram o mapeamento conforme das grades do plano s no plano $G(s)$. Quanto mais próximos os pólos de malha-fechada estiverem localizados do eixo $j\omega$, mais próximo o lugar geométrico de $G(j\omega)$ estará do ponto $-1 + j0$.

Suponha que o sistema possua pelo menos um par de pólos complexos conjugados de malha-fechada. Se o ponto $-1 - j0$ for determinado na intersecção de uma curva de σ constante e uma curva de ω constante no plano $G(s)$, então estes valores particulares de σ e ω, os quais definiremos como σ_c e ω_c, respectivamente, caracterizam o pólo de malha-fechada mais próximo do eixo $j\omega$ no semiplano superior do plano s. (Note que σ_c representa o decaimento exponencial e ω_c representa a

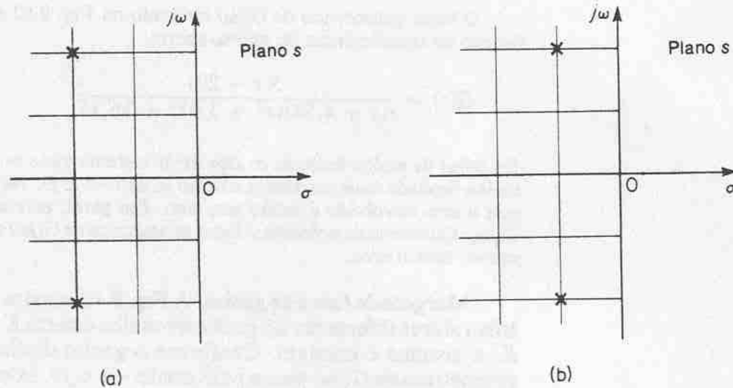


Fig. 9.58 Dois sistemas com dois pólos de malha-fechada.

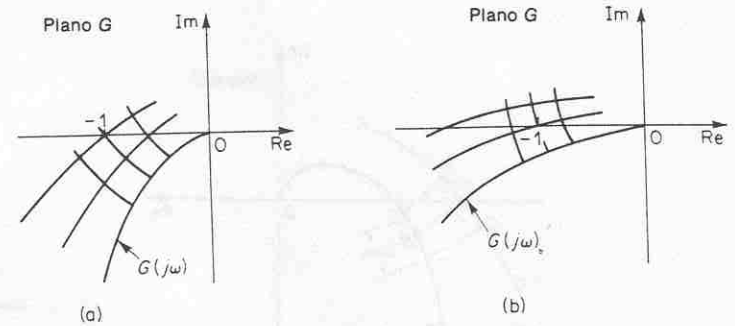


Fig. 9.59 Mapeamentos conformes de grades do plano s para os sistemas indicados na Fig. 9.58 para o plano $G(s)$.

freqüência natural amortecida do termo da resposta transitória ao degrau, devido ao par de pólos de malha-fechada mais próximo do eixo $j\omega$.) Se for conhecida a forma analítica de $G(s)$, o mapeamento conforme das retas $s = j\omega_1$, $s = j\omega_2$ e $s = -\sigma_1$ no plano $G(s)$ pode ser feito sem dificuldade. Então, os valores prováveis de σ_c e ω_c podem ser estimados a partir do gráfico, conforme indicado na Fig. 9.60. Portanto, o par de pólos complexos conjugados de malha-fechada que estão mais próximos do eixo $j\omega$ pode ser determinado graficamente. Deve ser observado que todos os pólos de malha-fechada estão mapeados no ponto $-1 + j0$ do plano $G(s)$. Embora os pólos complexos conjugados de malha-fechada mais próximos ao eixo $j\omega$ possam ser determinados facilmente pela técnica acima, a determinação de outros pólos de malha-fechada, se houver, é praticamente impossível por meio desta técnica.

Se os dados de $G(j\omega)$ forem experimentais, então pode ser construído por extrapolação um quadrado curvilíneo próximo ao ponto $-1 + j0$. Referindo-se à Fig. 9.61, podemos determinar a localização dos pólos de malha-fechada dominantes no plano s, ou a relação de amortecimento ζ e a freqüência natural amortecida ω_d, desenhando as retas AB que interligam o ponto $-1 + j0$ (ponto A) e o ponto B, a abordagem mais próxima do ponto $-1 + j0$, e então construindo um quadrado curvilíneo $CDEF$, como é indicado na Fig. 9.61. Este quadrado curvilíneo $CDEF$ pode ser construído desenhando-se a curva PQ mais parecida (onde PQ é o mapeamento conforme de uma reta paralela ao eixo $j\omega$ no plano s) passando pelo ponto $-1 + j0$ e "paralela" ao lugar geométrico de $G(j\omega)$, e ajustando os pontos C, D, E e F tais que $\widehat{FB} = \widehat{BE}$, $\widehat{CA} = \widehat{AD}$ e $\widehat{FE} + \widehat{CD} = \widehat{FC} + \widehat{ED}$. O contorno $C'D'E'F'$

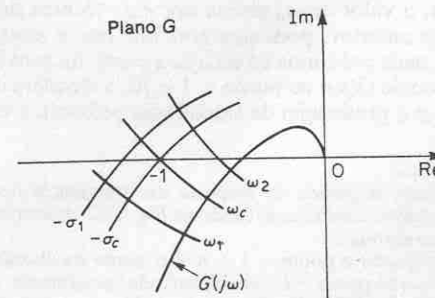


Fig. 9.60 Estimativa de σ_c e ω_c.

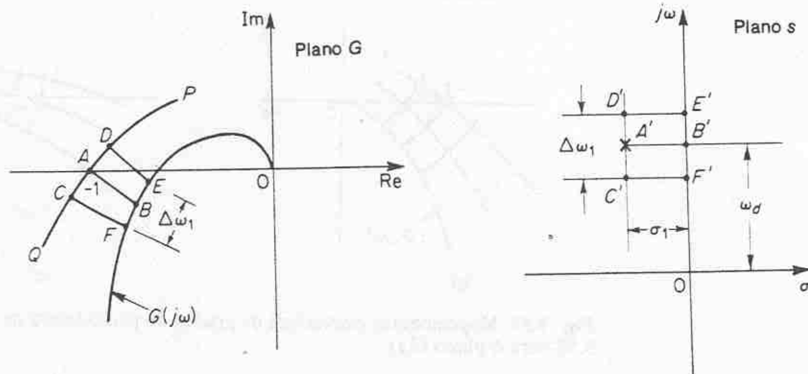


Fig. 9.61 Mapeamento conforme de um quadrado curvilíneo próximo ao ponto $-1 + j0$ no plano $G(s)$ para o plano s.

correspondente no plano s, junto com o ponto A', o pólo de malha-fechada mais próximo do eixo $j\omega$, são indicados na Fig. 9.61. O valor do intervalo de freqüência $\Delta\omega_1$ entre os pontos E e F é aproximadamente igual ao valor de σ_1 indicado na Fig. 9.61. A freqüência no ponto B é aproximadamente igual à freqüência natural amortecida ω_d. Os pólos de malha-fechada mais próximos ao eixo $j\omega$ são estimados como

$$s = -\sigma_1 \pm j\omega_d$$

Então, a relação de amortecimento ζ destes pólos de malha-fechada pode ser obtida de

$$\frac{\zeta}{\sqrt{1-\zeta^2}} = \frac{\sigma_1}{\omega_d} = \frac{\Delta\omega_1}{\omega_d}$$

Deve ser notado que a freqüência natural amortecida ω_d da resposta transitória ao degrau realmente está sobre o contorno de freqüência que passa pelo ponto $-1 + j0$ e não é necessariamente o ponto da abordagem mais próxima ao lugar geométrico. Portanto, o valor de ω_d obtida por esta técnica possui um certo erro.

Da análise anterior, podemos concluir que é possível estimar os pólos de malha-fechada mais próximos ao eixo $j\omega$ a partir da proximidade da abordagem do lugar geométrico de $G(j\omega)$ ao ponto $-1 + j0$, a freqüência no ponto da abordagem mais próxima, e a graduação de freqüência próxima a este ponto.

Exemplo 9.12 Dado o gráfico de resposta em freqüência de $G(j\omega)$ de um sistema com realimentação unitária, conforme indicado na Fig. 9.62, determine os pólos de malha-fechada mais próximos ao eixo $j\omega$.

A reta interligando o ponto $-1 + j0$ e o ponto da abordagem mais próxima do lugar geométrico de $G(j\omega)$ ao ponto $-1 + j0$ é desenhada inicialmente. Posteriormente, é construído o quadrado curvilíneo $ABCD$. Desde que a freqüência no ponto de abordagem mais próxima é $\omega = 2{,}9$, a freqüência natural amortecida é aproximadamente 2,9 ou $\omega_d = 2{,}9$. A partir do

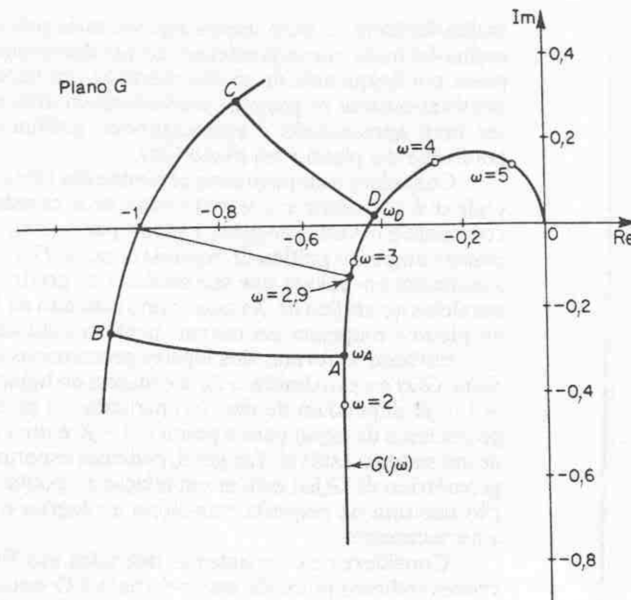


Fig. 9.62 Gráfico polar e um quadrado curvilíneo.

quadrado curvilíneo $ABCD$, verifica-se que

$$\Delta\omega = \omega_D - \omega_A = 3{,}4 - 2{,}4 = 1{,}0$$

Os pólos de malha-fechada mais próximos ao eixo $j\omega$ são então estimados como

$$s = -1 \pm j2{,}9$$

O lugar geométrico de $G(j\omega)$ indicado na Fig. 9.62 é realmente um gráfico da seguinte função de transferência de malha-aberta:

$$G(s) = \frac{5(s+20)}{s(s+4{,}59)(s^2+3{,}41s+16{,}35)}$$

Os pólos de malha-fechada exatos deste sistema são $s = -1 \pm j3$ e $s = -3 \pm j1$. Os pólos de malha-fechada mais próximos ao eixo $j\omega$ são $-1 \pm j3$. Neste exemplo particular, verificamos que o erro envolvido é muito pequeno. Em geral, este erro depende da particular curva de $G(j\omega)$. Quanto mais próximo o lugar geométrico de $G(j\omega)$ estiver em relação ao ponto $-1 + j0$, menor será o erro.

Margens de fase e de ganho. A Fig. 9.63 mostra os gráficos polares de $G(j\omega)$ para três valores diferentes do ganho de malha-aberta K. Para um valor grande do ganho K, o sistema é instável. Conforme o ganho diminui para um certo valor, o lugar geométrico de $G(j\omega)$ passa pelo ponto $-1 + j0$. Isto significa que, com este valor de ganho, o sistema está no limiar de instabilidade, e exibirá oscilações mantidas. Para um valor menor do ganho K, o sistema é estável.

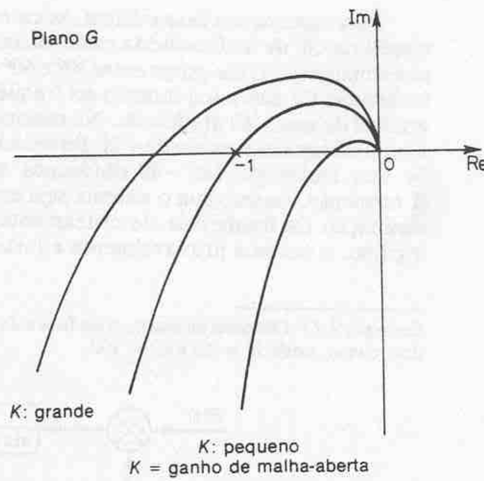


Fig. 9.63 Gráficos polares de $\dfrac{K(1+j\omega T_a)(1+j\omega T_b)\cdots}{(j\omega)^\lambda(1+j\omega T_1)(1+j\omega T_2)\cdots}$

Em geral, quanto mais próximo o lugar geométrico de $G(j\omega)$ resulta para envolver o ponto $-1+j0$, mais oscilatória é a resposta do sistema. A proximidade do lugar geométrico de $G(j\omega)$ ao ponto $-1+j0$ pode ser usada como uma medida da margem de estabilidade. (Isto não se aplica, entretanto, a sistemas condicionalmente estáveis.) Constitui uma prática comum representar a proximidade em termos da margem de fase e da margem de ganho.

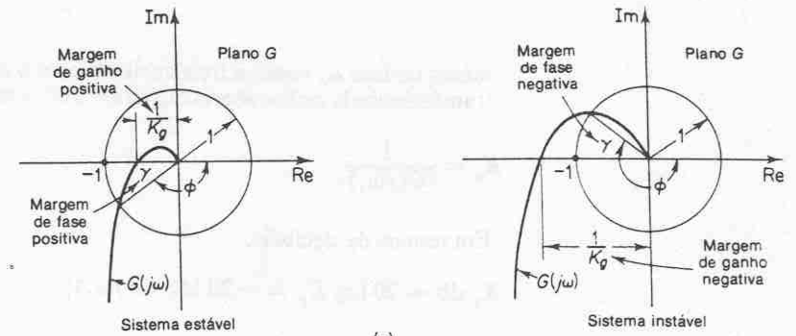
Margem de fase: A margem de fase é o atraso de fase adicional na freqüência de cruzamento do ganho, necessário para levar o sistema ao limiar de instabilidade. A freqüência de cruzamento do ganho é a freqüência na qual $|G(j\omega)|$, o módulo da função de transferência de malha-aberta, é unitário. A margem de fase γ é 180° mais o ângulo de fase ϕ da função de transferência de malha-aberta na freqüência de cruzamento do ganho, ou

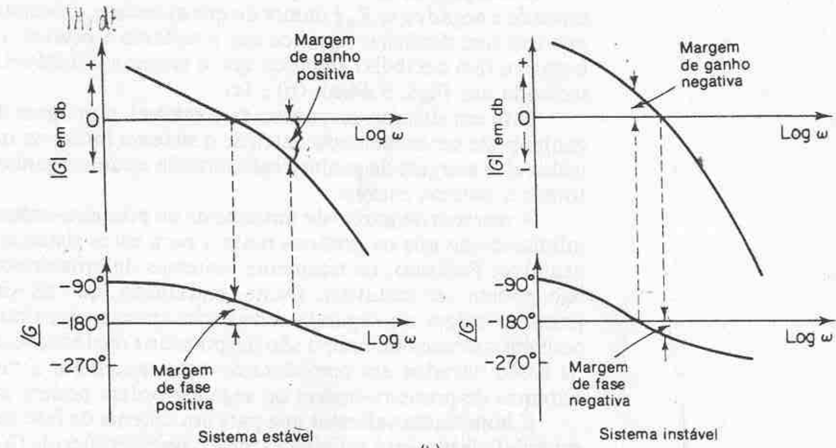
$$\gamma = 180^\circ + \phi$$

No diagrama de Nyquist, uma reta pode ser desenhada a partir da origem até o ponto no qual a circunferência de raio unitário intercepta o lugar geométrico de $G(j\omega)$. O ângulo a partir do eixo real negativo até esta reta é a margem de fase. A margem de fase é positiva para $\gamma > 0$ e negativa para $\gamma < 0$. Para um sistema de mínima fase ser estável, a margem de fase deve ser positiva.

As Figs. 9.64(a), (b) e (c) mostram as margens de fase de um sistema estável e de um sistema instável por meio de gráficos polares, gráficos logarítmicos e gráficos log-módulo *versus* fase. Nos gráficos logarítmicos, o ponto crítico no plano complexo corresponde às retas 0 db e −180°.

Margem de ganho: A margem de ganho é o recíproco do módulo $|G(j\omega)|$ na freqüência onde o ângulo de fase é −180°. Definindo a freqüência de cruza-





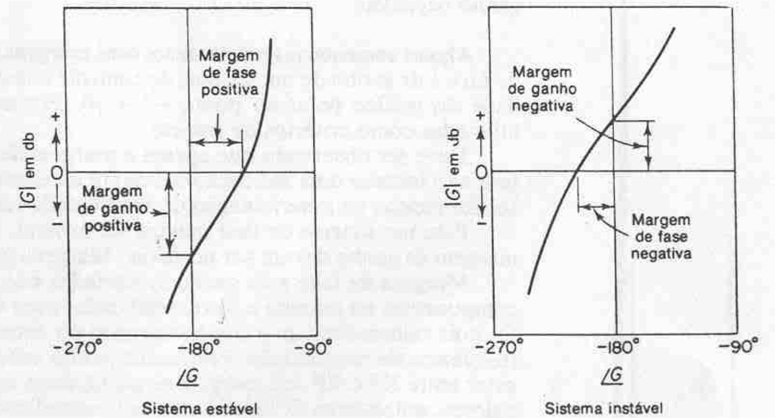


Fig. 9.64 Margens de ganho e de fase de sistemas estáveis e instáveis. (a) Gráficos polares; (b) gráficos logarítmicos; (c) gráficos do log-módulo *versus* fase.

mento de fase ω_1 como a freqüência na qual o ângulo de fase da função de transferência de malha-aberta é igual a $-180°$, a margem de ganho K_g resulta:

$$K_g = \frac{1}{|G(j\omega_1)|}$$

Em termos de decibéis,

$$K_g \text{ db} = 20 \log K_g = -20 \log |G(j\omega_1)|$$

A margem de ganho expressa em decibéis é positiva se K_g é maior do que a unidade e negativa se K_g é menor do que a unidade. Portanto, uma margem de ganho positiva (em decibéis) significa que o sistema é estável, e uma margem de ganho negativa (em decibéis) significa que o sistema é instável. A margem de ganho é indicada nas Figs. 9.64(a), (b) e (c).

Para um sistema de mínima fase estável, a margem de ganho indica quanto o ganho pode ser aumentado antes de o sistema tornar-se instável. Para um sistema instável, a margem de ganho é indicativa de quanto o ganho deve ser diminuído para tornar o sistema estável.

A margem de ganho de um sistema de primeira-ordem ou de segunda-ordem é infinita desde que os gráficos polares para estes sistemas não cruzem o eixo real negativo. Portanto, teoricamente, sistemas de primeira-ordem ou segunda-ordem não podem ser instáveis. (Note, entretanto, que os sistemas denominados de primeira-ordem ou segunda-ordem são apenas aproximações no sentido de que pequenos atrasos de tempo são desprezados na dedução das equações do sistema. Se forem levados em consideração estes pequenos atrasos, os assim chamados sistemas de primeira-ordem ou segunda-ordem podem tornar-se instáveis.)

É importante salientar que para um sistema de fase não-mínima, a condição de estabilidade não será satisfeita a menos que o gráfico de $G(j\omega)$ envolva o ponto $-1 + j0$. Portanto, um sistema de fase não-mínima estável possuirá margens de fase e de ganho negativas.

Alguns comentários relacionados com margens de fase e de ganho. As margens de fase e de ganho de um sistema de controle constituem uma medida da proximidade do gráfico polar ao ponto $-1 + j0$. Portanto, estas margens podem ser utilizadas como critérios de projeto.

Deve ser observado que apenas a margem de ganho ou apenas a margem de fase não fornece uma indicação suficiente da estabilidade relativa. Ambas devem ser fornecidas na determinação da estabilidade relativa.

Para um sistema de fase mínima ser estável, tanto a margem de fase como a margem de ganho devem ser positivas. Margens negativas indicam instabilidade.

Margens de fase e de ganho apropriadas nos previnem contra variações dos componentes no sistema e são especificadas para valores definidos de freqüência. Os dois valores limitam o comportamento do sistema de malha-fechada próximo à freqüência de ressonância. Para desempenho satisfatório, a margem de fase deve estar entre 30° e 60°, e a margem de ganho deve ser maior do que 6 db. Com estes valores, um sistema de fase mínima tem estabilidade garantida, mesmo se o ganho de malha-aberta e as constantes de tempo dos componentes variarem em uma grande extensão. Embora as margens de fase e de ganho forneçam apenas estimativas grosseiras da relação de amortecimento efetiva do sistema de malha-fechada, elas oferecem um meio conveniente no projeto de sistemas de controle ou no ajuste das constantes de ganho de sistemas.

Para sistemas de fase mínima, as características de módulo e fase da função de transferência de malha-aberta estão definitivamente relacionadas. A exigência de que a margem de fase esteja entre 30° e 60° significa que em um gráfico logarítmico a inclinação da curva log-módulo na freqüência de cruzamento do ganho seja mais gradual do que -40 db/década. Na maioria dos casos práticos, para estabilidade, é desejável uma inclinação de -20 db/década na freqüência de cruzamento do ganho. Se esta inclinação for -40 db/década o sistema pode ser estável ou instável. (Entretanto, mesmo que o sistema seja estável, a margem de fase é pequena.) Se a inclinação da freqüência de cruzamento do ganho for -60 db/década ou mais íngrime, o sistema provavelmente é instável.

Exemplo 9.13 Obtenha as margens de fase e de ganho do sistema indicado na Fig. 9.65 para os dois casos, onde $K = 10$ e $K = 100$.

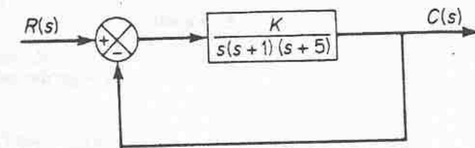


Fig. 9.65 Sistema de controle.

As margens de fase e de ganho podem ser facilmente obtidas a partir do gráfico logarítmico. Um gráfico logarítmico da função de transferência de malha-aberta com $K = 10$ é indicado na Fig. 9.66(a). As margens de fase e de ganho para $K = 10$ são

$$\text{Margem de fase} = 21°$$
$$\text{Margem de ganho} = 8 \text{ db}$$

Portanto, o ganho do sistema pode ser aumentado de 8 db antes de ocorrer instabilidade.

O aumento do ganho de $K = 10$ para $K = 100$ é correspondente a deslocar o eixo de 0-db para baixo por 20 db, conforme a Fig. 9.66(b). As margens de fase e de ganho são

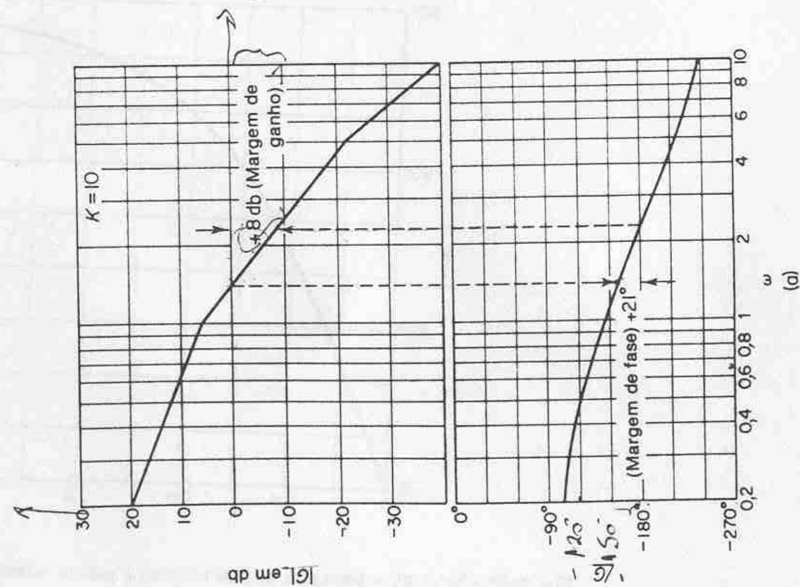
$$\text{Margem de fase} = -30°$$
$$\text{Margem de ganho} = -12 \text{ db}$$

Portanto, o sistema é estável para $K = 10$ porém instável para $K = 100$.

Note que para obter um desempenho satisfatório, devemos aumentar a margem de fase para 30° ~ 60°. Isto pode ser realizado diminuindo-se o ganho K. Entretanto, a diminuição de K não é desejável porque um pequeno valor de K resultará em um grande erro para a entrada em rampa. Isto sugere que pode ser necessária uma modificação na forma da curva de resposta em freqüência de malha-aberta, adicionando-se compensação. Tais técnicas são discutidas em detalhe no Cap. 10.

Correlação entre a resposta transitória ao degrau e a resposta em freqüência para sistemas de segunda-ordem. Para sistemas de segunda-ordem podem ser deduzidas as relações matemáticas exatas entre a resposta transitória ao degrau e a resposta em freqüência. Considere o sistema indicado na Fig. 9.67. A função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \qquad (9\text{-}15)$$



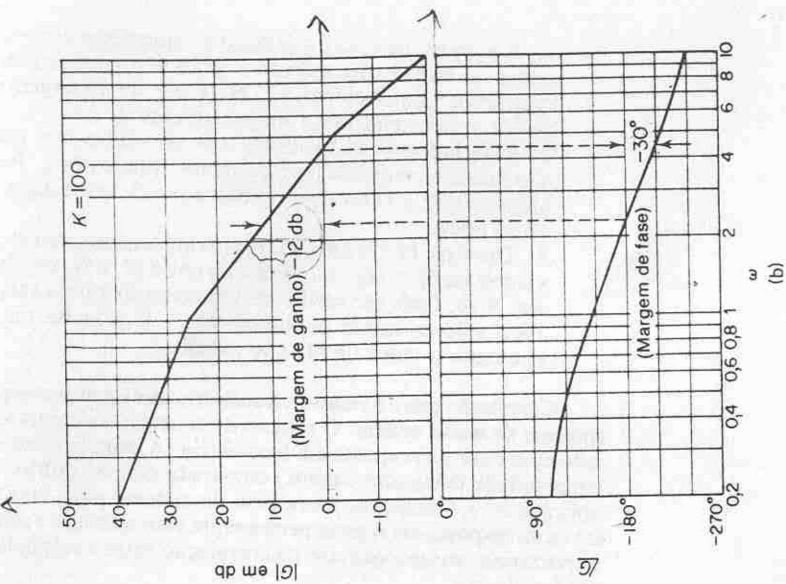


Fig. 9.66 Diagramas de Bode do sistema indicado na Fig. 9.65 com $K = 10$ e $K = 100$.

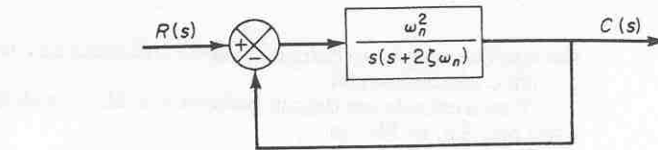


Fig. 9.67 Sistema de controle.

onde ζ e ω_n são a relação de amortecimento e a freqüência natural não amortecida, respectivamente. A resposta em freqüência de malha-fechada é

$$\frac{C(j\omega)}{R(j\omega)} = \frac{1}{\left(1 - \frac{\omega^2}{\omega_n^2}\right) + j2\zeta\frac{\omega}{\omega_n}} = Me^{j\alpha}$$

onde

$$M = \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_n^2}\right)^2 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2}}, \qquad \alpha = -\operatorname{tg}^{-1}\frac{2\zeta\frac{\omega}{\omega_n}}{1 - \frac{\omega^2}{\omega_n^2}}$$

Pela Eq. (9.12), para $0 \leqslant \zeta \leqslant 0{,}707$ o valor máximo de M ocorre na freqüência de ressonância ω_r, onde

$$\omega_r = \omega_n\sqrt{1 - 2\zeta^2} = \omega_n\sqrt{\cos 2\psi} \tag{9-16}$$

O ângulo ψ é definido na Fig. 9.68. Na freqüência de ressonância o valor de M é máximo e dado pela Eq. (9.13), resultando

$$M_r = \frac{1}{2\zeta\sqrt{1 - \zeta^2}} = \frac{1}{\operatorname{sen} 2\psi} \tag{9-17}$$

Lembre-se que ω_r é real somente se $\zeta < 0{,}707$. Portanto, não há ressonância de malha-fechada se $\zeta > 0{,}707$. [O valor de M_r é unitário para $\zeta > 0{,}707$. Vide a Eq. (9.14).] Desde que os valores de M_r e ω_r possam ser medidos facilmente em um

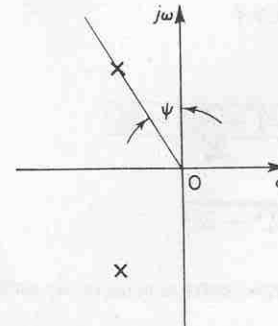


Fig. 9.68 Definição do ângulo ψ.

sistema físico, eles são bem úteis para verificação da concordância entre a análise teórica e a experimental.

Para a entrada em degrau unitário, a saída do sistema indicado na Fig. 9.67 é dada pela Eq. (6.19), ou

$$c(t) = 1 - e^{-\zeta\omega_n t}\left(\cos \omega_d t + \frac{\zeta}{\sqrt{1-\zeta^2}} \operatorname{sen} \omega_d t\right) \qquad (t \geq 0)$$

onde

$$\omega_d = \omega_n\sqrt{1-\zeta^2} = \omega_n \cos \psi \tag{9-18}$$

Por outro lado, a sobrelevação máxima M_p para a resposta ao degrau unitário é dada pela Eq. (6.28), ou

$$M_p = e^{-(\zeta/\sqrt{1-\zeta^2})\pi} \tag{9-19}$$

A sobrelevação máxima ocorre na resposta transitória que possui a freqüência natural amortecida $\omega_d = \omega_n\sqrt{1-\zeta^2}$. A máxima sobrelevação resulta excessiva para valores de $\zeta < 0,4$.

Desde que o sistema de segunda-ordem indicado na Fig. 9.67 possui a seguinte função de transferência de malha-aberta:

$$G(s) = \frac{\omega_n^2}{s(s + 2\zeta\omega_n)}$$

para operação senoidal, o módulo de $G(j\omega)$ resulta unitário quando

$$\omega = \omega_n\sqrt{\sqrt{1+4\zeta^4} - 2\zeta^2}$$

Nesta freqüência, o ângulo de fase de $G(j\omega)$ é

$$\underline{/G(j\omega)} = \underline{/\frac{1}{j\omega}} + \underline{/\frac{1}{j\omega + 2\zeta\omega_n}}$$
$$= -90° - \operatorname{tg}^{-1}\frac{\sqrt{\sqrt{1+4\zeta^4} - 2\zeta^2}}{2\zeta}$$

Portanto, a margem de fase γ é

$$\gamma = 180° + \underline{/G(j\omega)}$$
$$= 90° - \operatorname{tg}^{-1}\frac{\sqrt{\sqrt{1+4\zeta^4} - 2\zeta^2}}{2\zeta}$$
$$= \operatorname{tg}^{-1}\frac{2\zeta}{\sqrt{\sqrt{1+4\zeta^4} - 2\zeta^2}} \tag{9-20}$$

A Eq. (9.20) fornece a relação entre a relação de amortecimento ζ e a margem de fase γ.

A seguir, sintetizaremos a correlação entre a resposta transitória ao degrau e a resposta em freqüência para o sistema de segunda-ordem dado pela Eq. (9.15):

1. A margem de fase e a relação de amortecimento são diretamente relacionadas. A Fig. 9.69 fornece um gráfico da margem de fase γ em função da relação de amortecimento ζ. Note que uma margem de fase de 60° corresponde a uma relação de amortecimento de 0.6.
2. Referindo-se às Eqs. (9.16) e (9.18) verificamos que os valores de ω_r e ω_d são quase os mesmos para pequenos valores de ζ. Portanto, para pequenos valores de ζ, o valor de ω_r é indicativo da velocidade da resposta transitória do sistema.
3. Das Eqs. (9.17) e (9.19), vemos que quanto menor o valor de ζ, maiores os valores de M_r e M_p. A correlação entre M_r e M_p em função de ζ é indicada na Fig. 9.70. Pode ser verificada uma relação íntima entre M_r e M_p para $\zeta > 0.4$. Para valores muito pequenos de ζ, M_r resulta muito grande ($M_r \gg 1$), enquanto o valor de M_p não excede 1.

Correlação entre a resposta transitória ao degrau e a resposta em freqüência para sistemas de maior ordem. O projeto de sistemas de controle é muitas vezes conduzido com base na resposta em freqüência. A principal razão deste fato é a relativa simplicidade desta abordagem comparada com as outras. Desde que, em muitas aplicações, é a resposta transitória do sistema para entradas não periódicas, ao invés da resposta em regime permanente com entradas senoidais, que é de principal importância, surge a questão da correlação entre a resposta transitória e a resposta em freqüência.

Para o sistema de segunda-ordem indicado na Fig. 9.67 as relações matemáticas correlacionando a resposta transitória ao degrau e a resposta em freqüência podem ser obtidas facilmente. A resposta temporal de um sistema de segunda-ordem pode ser prevista exatamente a partir do conhecimento de M_r e ω_r de sua resposta em freqüência de malha-fechada.

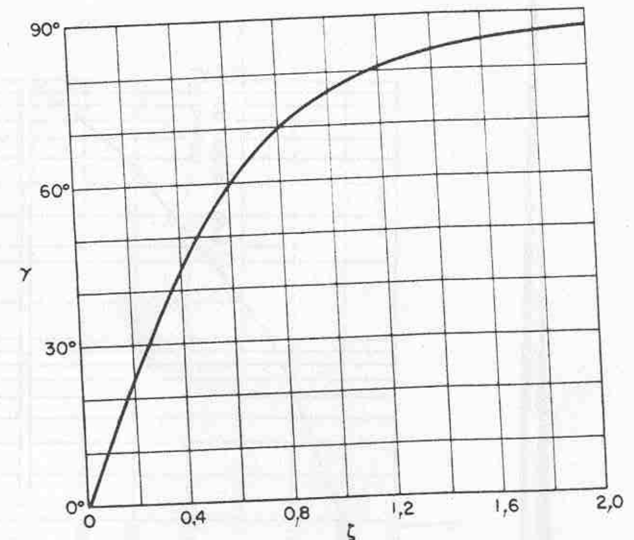


Fig. 9.69 Curva de γ (margem de fase) *versus* ζ para o sistema indicado na Fig. 9.67.

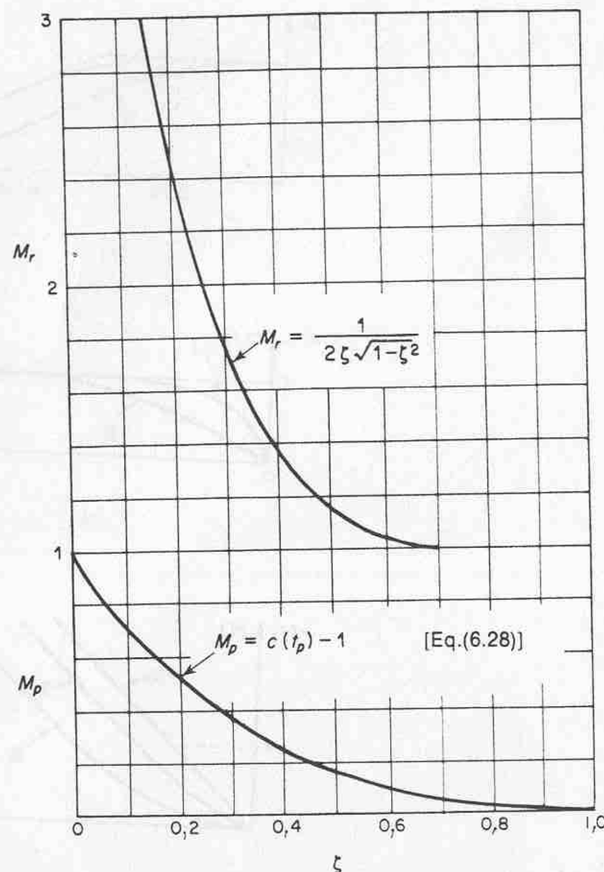


Fig. 9.70 Curvas de M_r *versus* ζ e M_p *versus* ζ para o sistema indicado na Fig. 9.67.

Para sistemas de maior ordem, a correlação é mais complexa e a resposta transitória pode não ser previsível facilmente a partir da resposta em freqüência, pois os pólos adicionais podem variar a correlação entre a resposta transitória ao degrau e a resposta em freqüência existente para um sistema de segunda-ordem. São disponíveis técnicas matemáticas para obtenção da correlação exata, porém são muito trabalhosas e de pequeno valor prático.

A aplicabilidade da correlação resposta-em-freqüência-resposta-transitória existente para o sistema de segunda-ordem indicado na Fig. 9.67 para sistemas de maior ordem depende da presença de um par dominante de pólos complexos conjugados de malha-fechada nestes últimos sistemas. É claro que, se a resposta em freqüência de um sistema de maior ordem é dominada por um par de pólos complexos conjugados de malha-fechada, a correlação resposta-em-freqüência-resposta-transitória existente para o sistema de segunda-ordem pode ser estendida para um sistema de maior ordem.

Para sistemas lineares, invariantes no tempo, e de maior ordem possuindo um par dominante de pólos complexos conjugados de malha-fechada, existem geralmente as seguintes relações entre a resposta transitória ao degrau e a resposta em freqüência:

1. O valor de M_r é indicativo da estabilidade relativa. Normalmente é obtido um desempenho transitório satisfatório se o valor de M_r estiver no intervalo $1.0 < M_r < 1.4$ ($0 \text{ db} < M_r < 3 \text{ db}$), que corresponde a uma relação de amortecimento efetiva de $0.4 < \zeta < 0.7$. Para valores de M_r maiores do que 1.5 a resposta transitória ao degrau pode exibir algumas sobrelevações. (Note que, em geral, um grande valor de M_r corresponde a uma grande sobrelevação na resposta transitória ao degrau. Se o sistema estiver sujeito a sinais com ruído cujas freqüências forem próximas à freqüência de ressonância ω_r, o ruído será amplificado na saída e causará um sério problema.)
2. O valor da freqüência de ressonância ω_r é indicativo da velocidade da resposta transitória. Quanto maior o valor de ω_r, mais rápido será o tempo de resposta. Em outras palavras, o tempo de subida varia inversamente com ω_r. Em termos da resposta em freqüência de malha-aberta, a freqüência natural amortecida da resposta transitória está entre a freqüência de cruzamento do ganho e a freqüência de cruzamento da fase.
3. A freqüência de ressonância ω_r e a freqüência natural amortecida ω_d para a resposta transitória ao degrau são muito próximas uma da outra para sistemas pouco amortecidos.

As três relações anteriormente alistadas são úteis para correlacionar a resposta transitória ao degrau com a resposta em freqüência de sistemas de maior ordem, desde que estes possam ser aproximados para um sistema de segunda-ordem ou a um par de pólos complexos conjugados de malha-fechada. Se o sistema de maior ordem satisfizer esta condição, um conjunto de especificações no domínio do tempo pode ser transferido para especificações no domínio de freqüência. Isto simplifica significativamente o trabalho de projeto, ou de compensação, de sistemas de maior ordem.

Além da margem de fase, margem de ganho, pico de ressonância M_r e freqüência de ressonância ω_r, há outras grandezas no domínio de freqüência comumente utilizadas em especificações e desempenho. Elas são a freqüência de corte, a largura de faixa e a taxa de corte. Estas grandezas serão definidas a seguir.

Freqüência de corte e largura de faixa. Referindo-se à Fig. 9.71, a freqüência ω_c na qual o módulo da resposta em freqüência de malha-fechada é 3 db abaixo de seu valor na freqüência-zero é denominada freqüência de corte. Portanto,

$$\left|\frac{C(j\omega)}{R(j\omega)}\right| < \left|\frac{C(j0)}{R(j0)}\right| - 3 \text{ db}$$
$$(\omega > \omega_c)$$

Para sistemas onde $|C(j0)/R(j0)| = 0$ db,

$$\left|\frac{C(j\omega)}{R(j\omega)}\right| < -3 \text{ db}$$
$$(\omega > \omega_c)$$

(Alguns autores usam o ponto de −6-db ao invés do ponto −3-db. Neste livro será utilizado o ponto −3-db para definir a freqüência de corte.) O sistema de malha-

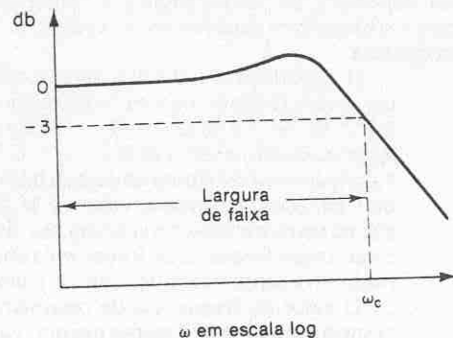


Fig. 9.71 Gráfico logarítmico indicando a freqüência de corte ω_c e a largura de faixa.

fechada filtra as componentes do sinal cujas freqüências são maiores do que a freqüência de corte e transmite aquelas componentes do sinal com freqüências menores do que a freqüência de corte.

A faixa de freqüência $0 \leq \omega \leq \omega_c$ na qual o módulo da malha fechada não cai -3 decibéis é denominada largura de faixa do sistema. A largura de faixa fornece uma indicação da velocidade da resposta de um sistema de controle.

A especificação da largura de faixa pode ser determinada pelos seguintes fatores:

1. A habilidade para reproduzir o sinal de entrada. (Uma largura de faixa grande corresponde a um pequeno tempo de subida, ou resposta rápida. Grosseiramente falando, podemos afirmar que a largura de faixa é proporcional à velocidade de resposta.)
2. As características de filtragem necessárias para o ruído de alta freqüência.

Para que o sistema siga precisamente entradas arbitrárias, é necessário que o sistema possua uma grande largura de faixa. Do ponto de vista do ruído, entretanto, a largura de faixa não deve ser demasiadamente grande. Portanto, há requisitos conflitantes em relação à largura de faixa e normalmente é necessário adotar-se um compromisso para um bom projeto. Note que um sistema com grande largura de faixa exige componentes com alto desempenho. (O custo dos componentes normalmente aumenta com a largura de faixa.)

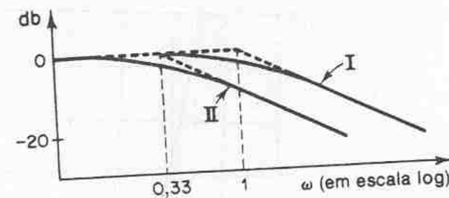
Taxa de corte. A taxa de corte é a inclinação da curva do log-módulo próxima à freqüência de corte. A taxa de corte indica a habilidade de um sistema para distinguir o sinal do ruído.

Exemplo 9.14 Considere os dois sistemas seguintes:
Sistema I:

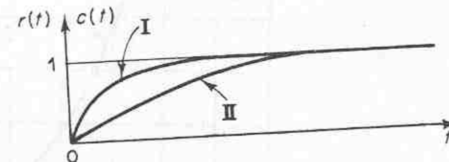
$$\frac{C(s)}{R(s)} = \frac{1}{s+1}$$

Sistema II:

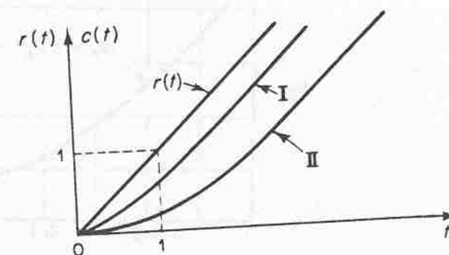
$$\frac{C(s)}{R(s)} = \frac{1}{3s+1}$$



(a)



(b)



(c)

Fig. 9.72 Comparação das características dinâmicas dos dois sistemas considerados no Exemplo 9.14.

Compare as larguras de faixas destes dois sistemas. Mostre que o sistema com a largura de faixa maior possui uma velocidade de resposta maior e pode seguir a entrada muito melhor do que aquele com uma largura de faixa menor.

A Fig. 9.72(a) indica as curvas de resposta em freqüência de malha-fechada para os dois sistemas. (Curvas assintóticas são indicadas por semi-retas tracejadas.) Verificamos que a largura de faixa do sistema I é $0 \leq \omega \leq 1$ rad/s e aquele do sistema II é $0 \leq \omega \leq 0,33$ rad/s. As Figs. 9.72(b) e (c) mostram, respectivamente, as curvas de resposta ao degrau e resposta à rampa para os dois sistemas. Obviamente, o sistema I, cuja largura de faixa é três vezes maior que a do sistema II, possui uma velocidade de resposta mais rápida e pode seguir melhor a entrada.

9.8 RESPOSTA EM FREQÜÊNCIA DE MALHA-FECHADA

Resposta em freqüência de malha-fechada de sistemas com realimentação unitária. Para um sistema de malha-fechada estável, a resposta em freqüência pode ser

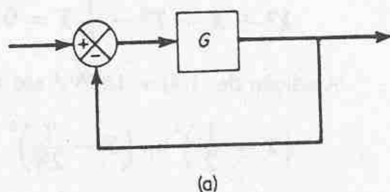
facilmente obtida a partir da resposta em malha-aberta. Considere o sistema com realimentação unitária indicado na Fig. 9.73(a). A função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)}$$

No gráfico de Nyquist ou polar indicado na Fig. 9.73(b), o vetor $\overrightarrow{OA}$ representa $G(j\omega_1)$ onde ω_1 é a freqüência no ponto A. O comprimento do vetor $\overrightarrow{OA}$ é $|G(j\omega_1)|$ e o ângulo do vetor $\overrightarrow{OA}$ é $\underline{/G(j\omega_1)}$. O vetor $\overrightarrow{PA}$, o vetor do ponto $-1 + j0$ para o lugar geométrico de Nyquist, representa $1 + G(j\omega_1)$. Portanto, a relação entre $\overrightarrow{OA}$ para $\overrightarrow{PA}$ representa a resposta em freqüência de malha-fechada, ou

$$\frac{\overrightarrow{OA}}{\overrightarrow{PA}} = \frac{G(j\omega_1)}{1 + G(j\omega_1)} = \frac{C(j\omega_1)}{R(j\omega_1)}$$

O módulo da função de transferência de malha-fechada em $\omega = \omega_1$ é a relação entre os módulos de $\overrightarrow{OA}$ para $\overrightarrow{PA}$. O ângulo de fase da função de transferência de malha-fechada em $\omega = \omega_1$ é o ângulo formado pelos vetores $\overrightarrow{OA}$ e $\overrightarrow{PA}$, isto é, $\phi - \theta$, indicado na Fig. 9.73(b). Medindo-se o módulo e o ângulo de fase em pontos de



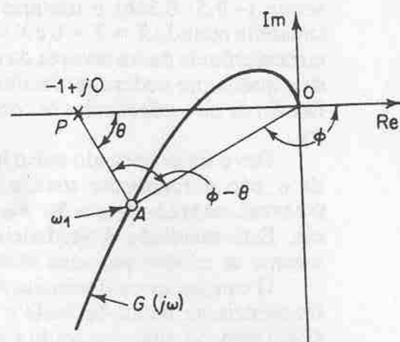


Fig. 9.73 (a) Sistema com realimentação unitária; (b) determinação da resposta em freqüência de malha-fechada a partir da resposta em freqüência de malha-aberta.

diferentes freqüências, pode ser obtida a curva de resposta em freqüência em malha-fechada.

Vamos definir o módulo da resposta em freqüência de malha-fechada como M e o ângulo de fase como α, ou

$$\frac{C(j\omega)}{R(j\omega)} = Me^{j\alpha}$$

A seguir, determinaremos os lugares geométricos do módulo constante e os lugares geométricos de ângulo de fase constante. Tais lugares geométricos são convenientes na determinação da resposta em freqüência de malha-fechada a partir do gráfico de Nyquist.

Lugares geométricos de módulo constante (circunferências *M*). Para obter os lugares geométricos de módulo constante, observamos inicialmente que $G(j\omega)$ é uma grandeza complexa e pode ser escrita como segue:

$$G(j\omega) = X + jY$$

onde X e Y são grandezas reais. Então M é dado por

$$M = \frac{|X + jY|}{|1 + X + jY|}$$

e M^2 é

$$M^2 = \frac{X^2 + Y^2}{(1 + X)^2 + Y^2}$$

Portanto

$$X^2(1 - M^2) - 2M^2X - M^2 + (1 - M^2)Y^2 = 0 \qquad (9\text{-}21)$$

Se $M = 1$, então, da Eq. (9.21) obtemos $X = -1/2$. Esta é a equação de uma reta paralela ao eixo Y e passando pelo ponto $(-1/2; 0)$.

Se $M \neq 1$, a Eq. (9.21) pode ser escrita

$$X^2 + \frac{2M^2}{M^2 - 1}X + \frac{M^2}{M^2 - 1} + Y^2 = 0$$

Se o termo $M^2/(M^2 - 1)^2$ for adicionado aos dois membros desta última equação, obtemos

$$\left(X + \frac{M^2}{M^2 - 1}\right)^2 + Y^2 = \frac{M^2}{(M^2 - 1)^2} \qquad (9\text{-}22)$$

A Eq. (9.22) é a equação de uma circunferência com centro em $X = -M^2/(M^2 - 1)$, $Y = 0$ e com raio $|M/(M^2 - 1)|$.

Os lugares geométricos M constante no plano $G(s)$ constituem, portanto, uma família de circunferências. O centro e o raio da circunferência para um dado valor de M podem ser calculados facilmente. Por exemplo, para $M = 1{,}3$, o centro está em $(-2{,}45; 0)$ e o raio é $1{,}88$. Uma família de circunferências M constante é indicada na Fig. 9.74. Pode ser visto que conforme M aumenta e torna-se maior do que a

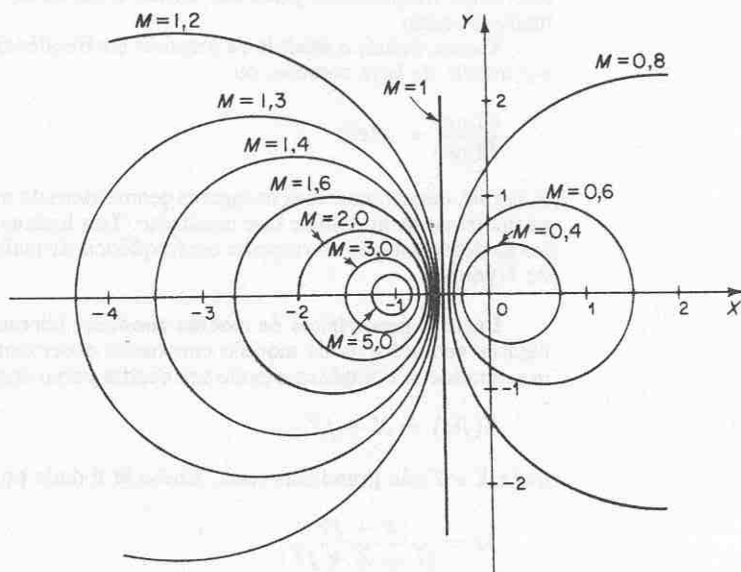


Fig. 9.74 Uma família de circunferências de M constante.

unidade, as circunferências tornam-se cada vez menores e convergem para o ponto $-1 + j0$. Para $M > 1$, os centros das circunferências M estão à esquerda do ponto $-1 + j0$. Analogamente, conforme M diminui e torna-se muito menor do que a unidade a circunferência M fica cada vez menor e converge para a origem. Para $M < 1$ os centros das circunferências M estão à direita da origem. $M = 1$ corresponde ao lugar geométrico dos pontos eqüidistantes da origem e do ponto $-1 + j0$. Conforme estabelecido anteriormente, este é uma reta passando pelo ponto $(-1/2, 0)$ e paralela ao eixo imaginário. (As circunferências M constante correspondentes a $M > 1$ estão à esquerda da reta $M = 1$, e aquelas correspondentes a $M < 1$ estão à direita da reta $M = 1$.) As circunferências são simétricas em relação à reta correspondente a $M = 1$ e em relação ao eixo real.

Lugares geométricos de ângulos de fase constante (circunferências N). Obteremos inicialmente o ângulo de fase α em termos de X e Y. Desde que

$$\underline{/e^{j\alpha}} = \underline{/\frac{X + jY}{1 + X + jY}}$$

o ângulo de fase α é

$$\alpha = \text{tg}^{-1}\left(\frac{Y}{X}\right) - \text{tg}^{-1}\left(\frac{Y}{1 + X}\right)$$

Se definirmos

$$\text{tg } \alpha = N$$

então

$$N = \text{tg}\left[\text{tg}^{-1}\left(\frac{Y}{X}\right) - \text{tg}^{-1}\left(\frac{Y}{1 + X}\right)\right]$$

Desde que

$$\text{tg } (A - B) = \frac{\text{tg } A - \text{tg } B}{1 + \text{tg } A \text{ tg } B}$$

obtemos

$$N = \frac{\dfrac{Y}{X} - \dfrac{Y}{1 + X}}{1 + \dfrac{Y}{X}\left(\dfrac{Y}{1 + X}\right)}$$

$$= \frac{Y}{X^2 + X + Y^2}$$

ou

$$X^2 + X + Y^2 - \frac{1}{N}Y = 0$$

A adição de $(1/4) + 1/(2N)^2$ aos dois membros desta última equação resulta em

$$\left(X + \frac{1}{2}\right)^2 + \left(Y - \frac{1}{2N}\right)^2 = \frac{1}{4} + \left(\frac{1}{2N}\right)^2 \qquad (9\text{-}23)$$

Esta é a equação de uma circunferência com centro em $X = -1/2$, $Y = 1/(2N)$ e com raio $\sqrt{(1/4) + 1/(2N)^2}$. Por exemplo, se $\alpha = 30°$, então $N = \text{tg}\,\alpha = 0{,}577$, e o centro e o raio da circunferência correspondente a $\alpha = 30°$ podem ser determinados como sendo $(-0{,}5;\ 0{,}866)$ e unitário, respectivamente. Desde que a Eq. (9.23) está satisfeita quando $X = Y = 0$ e $X = -1$, $Y = 0$ independentemente do valor de N, cada circunferência passa através da origem e do ponto $-1 + j0$. Os lugares geométricos de α constante podem ser facilmente desenhados uma vez dado o valor de N. Uma família de circunferências N constante é indicada na Fig. 9.75 com α como parâmetro.

Deve ser observado que o lugar geométrico de N constante para um dado valor de α não é realmente toda a circunferência mas apenas um arco. Em outras palavras, os arcos de $\alpha = 30°$ e $\alpha = -150°$ constituem partes da mesma circunferência. Este resultado é verdadeiro porque a tangente de um ângulo permanece a mesma se $\pm$ 180° (ou seus múltiplos) forem adicionados ao ângulo.

O uso das circunferências M e N nos possibilita determinar toda a resposta em freqüência de malha-fechada a partir da resposta em freqüência de malha-aberta $G(j\omega)$ sem calcular o módulo e a fase da função de transferência de malha-fechada em cada freqüência. As intersecções do lugar geométrico de $G(j\omega)$ e as circunferências M e as circunferências N fornecem os valores de M e N nos pontos de freqüência sobre o lugar geométrico de $G(j\omega)$.

Desde que as circunferências N são de múltiplos valores, no sentido que a circunferência para $\alpha = \alpha_1$ e uma outra para $\alpha = \alpha_1 \pm 180°n$ $(n = 1, 2, \ldots)$ na realidade são as mesmas, usando-se as circunferências N para a determinação do

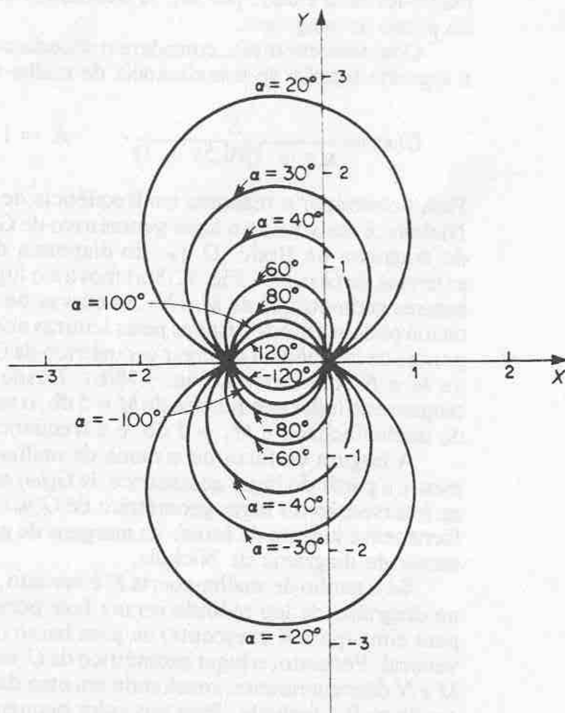


Fig. 9.75 Uma família de circunferências com N constante.

ângulo de fase de sistemas de malha-fechada, devemos interpretar o valor apropriado de α. Para evitar qualquer erro, inicia-se na freqüência zero, que corresponde a $\alpha = 0°$, e então efetua-se o procedimento para freqüências maiores. A curva do ângulo de fase deve ser contínua.

A Fig. 9.76(a) mostra o lugar geométrico de $G(j\omega)$ sobreposto a uma família de circunferências M. A Fig. 9.76(b) mostra o lugar geométrico de $G(j\omega)$ sobreposto a uma família de circunferências N. Destes gráficos, é possível obter por inspeção a resposta em freqüência de malha-fechada. Verifica-se que a circunferência $M = 1{,}1$ intercepta o lugar geométrico de $G(j\omega)$ no ponto de freqüência $\omega = \omega_1$. Isto significa que nesta freqüência o módulo da freqüência de malha-fechada é 1,1. Na Fig. 9.76(a), a circunferência $M = 2$ é exatamente tangente ao lugar geométrico de $G(j\omega)$. Portanto, há apenas um ponto no lugar geométrico de $G(j\omega)$ para o qual $|C(j\omega)|R(j\omega)|$ é igual a 2. A Fig. 9.76(c) mostra a curva de resposta em freqüência de malha-fechada para o sistema. A curva superior é a curva M *versus* freqüência ω e a curva inferior é a curva ângulo de fase α *versus* freqüência ω.

O valor de pico de ressonância é o valor de M correspondente à circunferência M de mínimo raio que é tangente ao lugar geométrico de $G(j\omega)$. Portanto, no diagrama de Nyquist, o valor do pico de ressonância M_r e a freqüência de ressonância ω_r podem ser determinados a partir da circunferência M que tangencia o lugar geométrico de $G(j\omega)$. (No presente exemplo, $M_r = 2$ e $\omega_r = \omega_4$.)

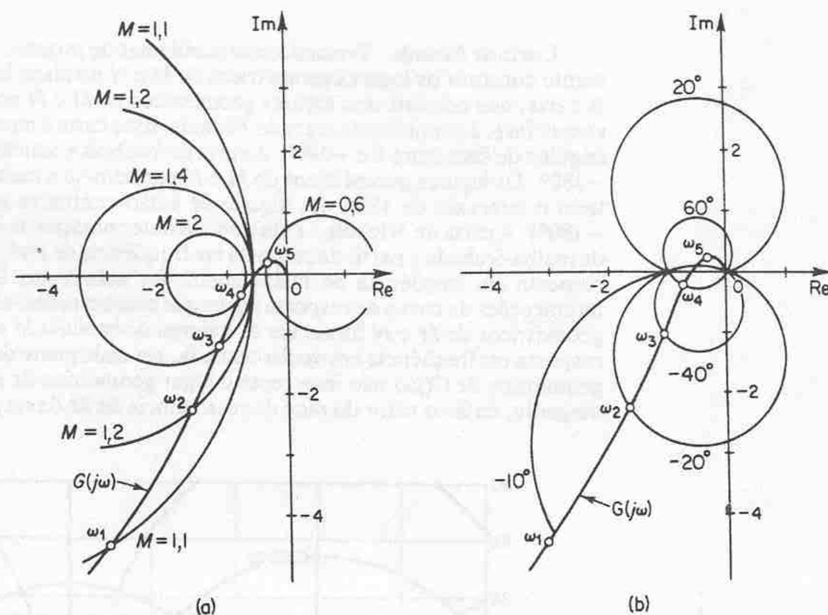


Fig. 9.76 (a) Lugar geométrico de $G(j\omega)$ sobreposto a uma família de circunferências M; (b) lugar geométrico de $G(j\omega)$ sobreposto a uma família de circunferências N; (c) curvas de resposta em freqüência de malha-fechada.

Carta de Nichols. Tratando com problemas de projeto, verificamos ser conveniente construir os lugares geométricos de M e N no plano log-módulo *versus* fase. A carta, que consiste nos lugares geométricos de M e N no diagrama log-módulo *versus* fase, é denominada carta de Nichols. Esta carta é mostrada na Fig. 9.77 para ângulos de fase entre 0 e $-240°$. A carta de Nichols é simétrica em relação ao eixo $-180°$. Os lugares geométricos de M e N repetem-se a cada 360° e há simetria em todo o intervalo de 180°. Os lugares M estão centrados no ponto crítico (0 db; $-180°$). A carta de Nichols é muito útil na determinação da resposta em freqüência de malha-fechada a partir da resposta em freqüência de malha-aberta. Se a curva de resposta em freqüência de malha-aberta for sobreposta à carta de Nichols, as intersecções da curva de resposta em freqüência em malha-aberta $G(j\omega)$ e os lugares geométricos de M e N fornecem os valores do módulo M e o ângulo de fase α da resposta em freqüência em malha-fechada, em cada ponto de freqüência. Se o lugar geométrico de $G(j\omega)$ não intercepta o lugar geométrico de $M = M_r$, porém é a ele tangente, então o valor do pico de ressonância de M da resposta em freqüência de

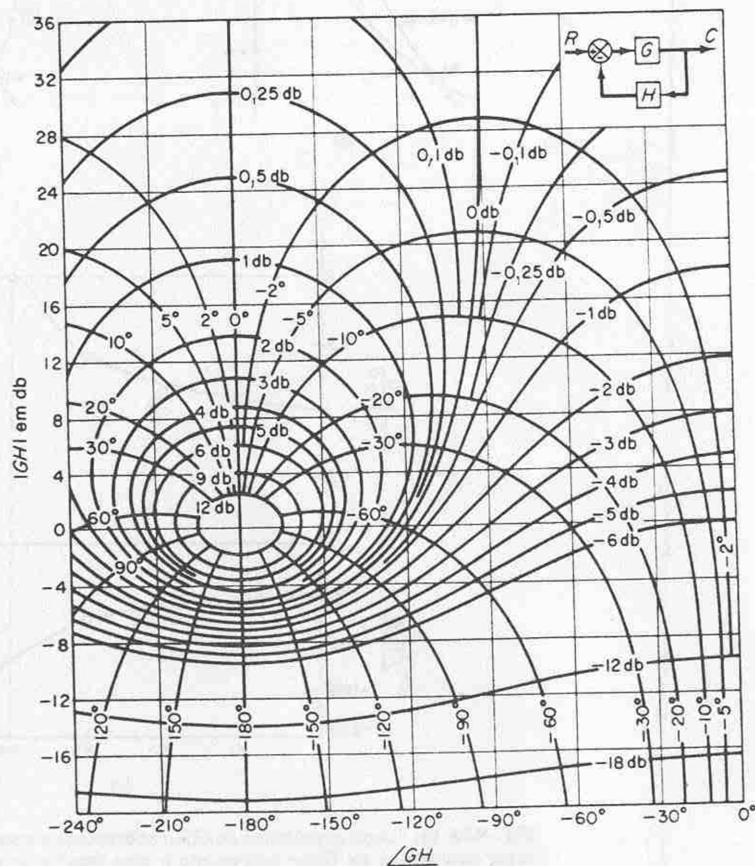


Fig. 9.77 Carta de Nichols.

malha-fechada é dado por M_r. A freqüência de ressonância é dada pela freqüência no ponto de tangência.

Como um exemplo, considere o sistema com realimentação unitária possuindo a seguinte função de transferência de malha-aberta:

$$G(s) = \frac{K}{s(s+1)(0{,}5s+1)}, \qquad K = 1$$

Para determinar a resposta em freqüência de malha-fechada utilizando a carta de Nichols, é construído o lugar geométrico de $G(j\omega)$ no plano log-módulo *versus* fase do diagrama de Bode. O uso do diagrama de Bode elimina cálculos numéricos extensos de $G(j\omega)$. A Fig. 9.78(a) mostra o lugar geométrico de $G(j\omega)$ junto com os lugares geométricos de M e N. As curvas de resposta em freqüência de malha-fechada podem ser construídas pelas leituras dos módulos e ângulos de fase em vários pontos de freqüência do lugar geométrico de $G(j\omega)$ a partir dos lugares geométricos de M e N, conforme a Fig. 9.78(b). Desde que o lugar geométrico de $G(j\omega)$ é tangente ao lugar geométrico de $M = 5$ db, o valor de pico da resposta em freqüência de malha-fechada é $M_r = 5$ db, e a freqüência de ressonância é 0,8 rad/s.

A largura de faixa do sistema de malha-fechada pode ser determinada facilmente a partir do lugar geométrico de $G(j\omega)$ no diagrama de Nichols. A freqüência na intersecção do lugar geométrico de $G(j\omega)$ e o lugar geométrico de $M = -3$ db fornecem a largura de faixa. As margens de ganho e de fase podem ser lidas diretamente do diagrama de Nichols.

Se o ganho de malha-aberta K é variado, a forma do lugar geométrico de $G(j\omega)$ no diagrama de log-módulo *versus* fase permanece a mesma, porém é deslocada para cima (para K crescente) ou para baixo (para K decrescente) ao longo do eixo vertical. Portanto, o lugar geométrico de $G(j\omega)$ intercepta os lugares geométricos de M e N diferentemente, resultando em uma diferente curva de resposta em freqüência de malha-fechada. Para um valor pequeno do ganho K, o lugar geométrico de $G(j\omega)$ não será tangente a qualquer lugar geométrico M, significando que não há ressonância na resposta em freqüência de malha-fechada.

Resposta em freqüência de malha-fechada para sistemas com realimentação não unitária. Se o sistema de malha-fechada envolver uma função de transferência com realimentação não unitária, então a função de transferência de malha-fechada pode ser escrita

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)}$$

onde $G(s)$ é a função de transferência do ramo direto e $H(s)$ é a função de transferência do ramo de realimentação. Então, $C(j\omega)/R(j\omega)$ pode ser escrita

$$\frac{C(j\omega)}{R(j\omega)} = \frac{1}{H(j\omega)}\,\frac{G(j\omega)H(j\omega)}{1 + G(j\omega)H(j\omega)}$$

O módulo e o ângulo de fase de

$$\frac{G_1(j\omega)}{1 + G_1(j\omega)}$$

onde $G_1(j\omega) = G(j\omega)H(j\omega)$, podem ser facilmente obtidos construindo-se o lugar geométrico de $G_1(j\omega)$ sobre a carta de Nichols e lendo os valores de M e N em vários

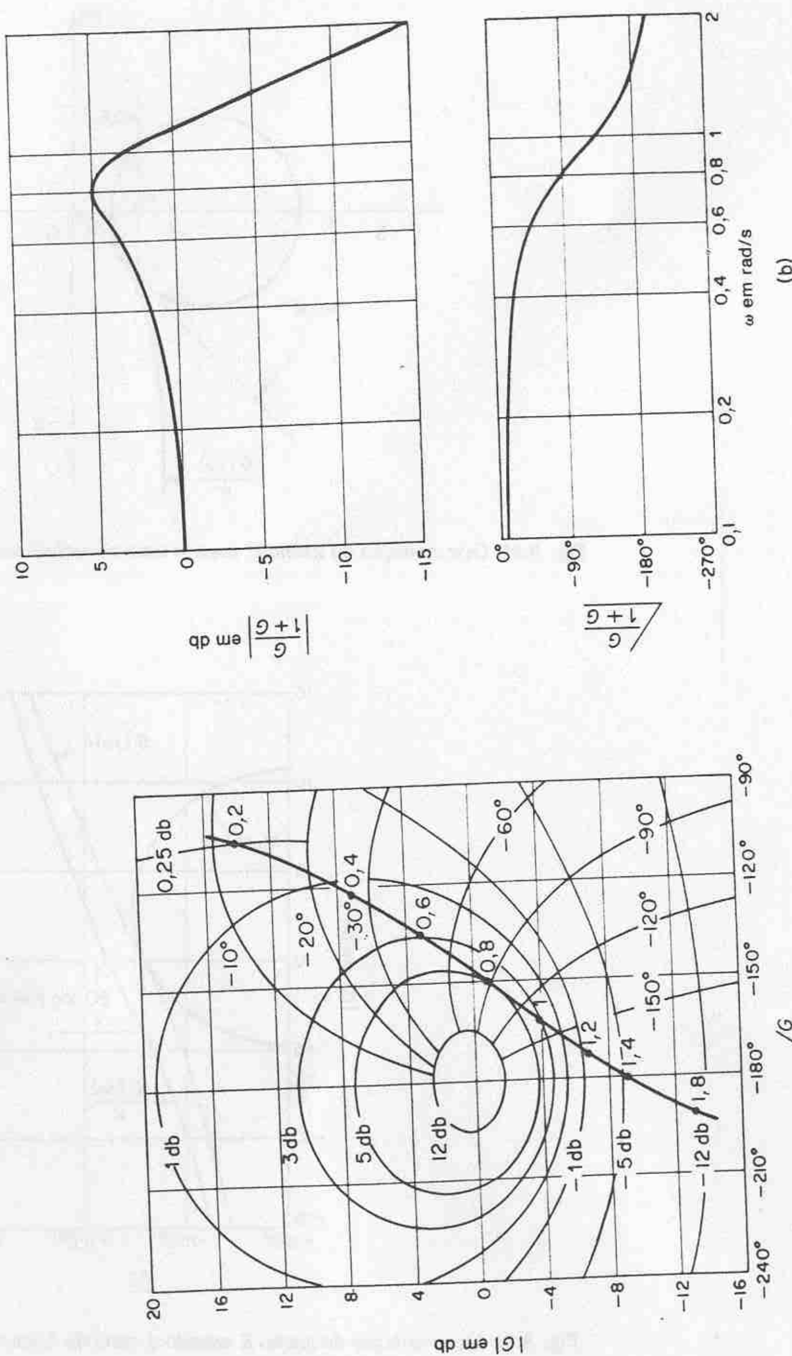


Fig. 9.78 (a) Gráfico de $G(j\omega)$ sobreposto à carta de Nichols; (b) curvas de resposta em freqüência de malha-fechada.

pontos de freqüência. A resposta em freqüência de malha-fechada $C(j\omega)/R(j\omega)$ pode ser então obtida multiplicando-se $G_1(j\omega)/[1 + G_1(j\omega)]$ por $1/H(j\omega)$. Esta multiplicação pode ser realizada sem dificuldade se desenharmos os diagramas de Bode para $G_1(j\omega)/[1 + G_1(j\omega)]$ e $H(j\omega)$ e então subtrairmos graficamente o módulo de $H(j\omega)$ daquele de $G_1(j\omega)/[1 + G_1(j\omega)]$ e também subtrairmos graficamente o ângulo de fase de $H(j\omega)$ daquele de $G_1(j\omega)/[1 + G_1(j\omega)]$. A curva do log-módulo e a curva do ângulo de fase resultantes fornecem, então, a resposta em freqüência de malha-fechada $C(j\omega)/R(j\omega)$.

Para obter valores aceitáveis de M_r, ω_r, e ω_c para $|C(j\omega)/R(j\omega)|$, pode ser necessário um processo de tentativa-e-erro. Em cada tentativa, o lugar geométrico de $G_1(j\omega)$ é variado em sua forma. São então desenhados os diagramas de Bode para $G_1(j\omega)/[1 + G_1(j\omega)]$ e $H(j\omega)$ e a resposta em freqüência de malha-fechada $C(j\omega)/R(j\omega)$ é obtida. Os valores de M_r, ω_r e ω_c são verificados até obterem-se valores aceitáveis.

Ajustes de ganho. O conceito das circunferências M será agora aplicado ao projeto de sistemas de controle. Para obter-se desempenho conveniente, o ajuste do ganho normalmente é a primeira consideração. O ajuste pode ser baseado em um valor desejado para o pico de ressonância.

Em seguida, demonstraremos o método para determinação do ganho K de modo que o sistema possuirá algum valor M_r máximo, não excedido em toda a faixa de freqüência.

Referindo-se à Fig. 9.79, verificamos que a reta tangente desenhada a partir da origem até a circunferência M_r desejada possui um ângulo ψ como indicado, se M_r for maior do que a unidade. O valor de sen ψ é

$$\text{sen}\,\psi = \frac{\dfrac{M_r}{M_r^2-1}}{\dfrac{M_r^2}{M_r^2-1}} = \frac{1}{M_r}$$

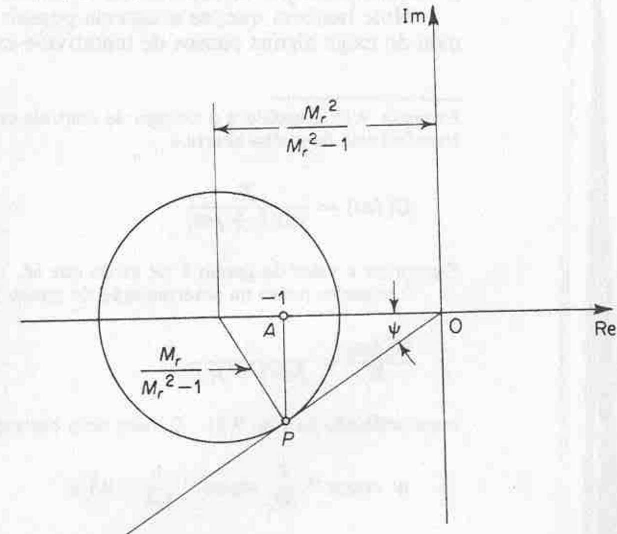


Fig. 9.79 Circunferência M.

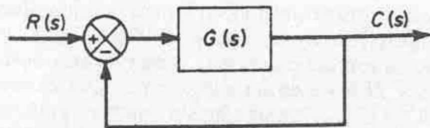


Fig. 9.80 Sistema de controle.

Pode-se provar facilmente que a reta desenhada pelo ponto P, perpendicular ao eixo real negativo, intercepta este eixo no ponto $-1 + j0$.

Considere o sistema indicado na Fig. 9.80. O procedimento para determinação do ganho K, de modo que $G(j\omega) = KG_1(j\omega)$ possuirá um valor desejado de M_r (onde $M_r > 1$), pode ser sumariado como segue:

1. Desenhar o gráfico polar da função de transferência de malha-aberta normalizada.
2. Desenhar a partir da origem a reta que faz com o eixo real negativo um ângulo de $\psi = \text{sen}^{-1}(1/M_r)$.
3. Desenhar uma circunferência com centro sobre o eixo real negativo e tangente tanto ao lugar geométrico de $G_1(j\omega)$ como à reta PO.
4. Desenhar uma reta perpendicular ao eixo real negativo pelo ponto P, o ponto de tangência desta circunferência com a reta PO. A reta perpendicular PA intercepta o eixo real negativo no ponto A.
5. Para que a circunferência determinada corresponda à circunferência M_r desejada, o ponto A deve ser o ponto $-1 + j0$.
6. O valor desejado do ganho K é aquele valor que modifica a escala de modo que o ponto A se torne o ponto $-1 + j0$. Portanto, $K = 1/\overline{OA}$.

Note que a freqüência de ressonância ω_r é a freqüência do ponto no qual a circunferência é tangente ao lugar geométrico de $G_1(j\omega)$. O procedimento apresentado pode não fornecer um valor satisfatório para ω_r. Se este for o caso, o sistema deve ser compensado a fim de aumentar o valor de ω_r sem modificar o valor de M_r. (Para compensação de sistemas de controle, vide Cap. 10.)

Note também que, se o sistema possuir realimentação não unitária, então o método exige alguns passos de tentativa-e-erro.

Exemplo 9.15 Considere o sistema de controle com realimentação unitária cuja função de transferência de malha-aberta é

$$G(j\omega) = \frac{K}{j\omega(1 + j\omega)}$$

Determine o valor do ganho K de modo que $M_r = 1{,}4$.

O primeiro passo na determinação do ganho K é esboçar o gráfico polar de

$$\frac{G(j\omega)}{K} = \frac{1}{j\omega(1 + j\omega)}$$

como indicado na Fig. 9.81. O valor de ψ correspondente a $M_r = 1{,}4$ é obtido de

$$\psi = \text{sen}^{-1}\frac{1}{M_r} = \text{sen}^{-1}\frac{1}{1{,}4} = 45{,}6°$$

O próximo passo é desenhar a reta OP que faz um ângulo de $\psi = 45{,}6°$ com o eixo real negativo. Desenhe, então, a circunferência tangente tanto ao lugar geométrico $G(j\omega)/K$ como

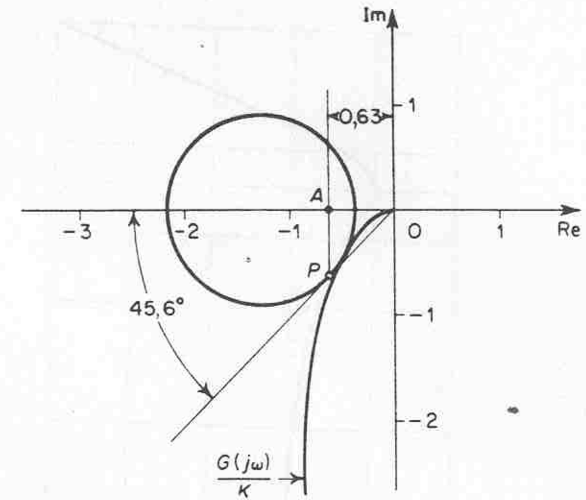


Fig. 9.81 Determinação do ganho K usando uma circunferência M.

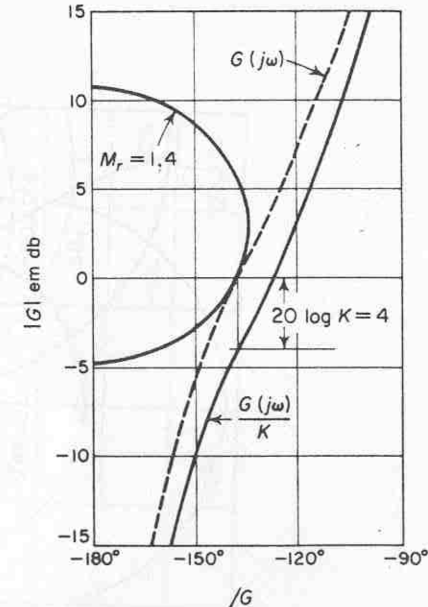


Fig. 9.82 Determinação do ganho K usando a carta de Nichols.

à reta OP. A reta perpendicular desenhada pelo ponto P intercepta o eixo real negativo em $(-0,63;\ 0)$. Então, o ganho K do sistema é determinado como segue:

$$K = \frac{1}{0,63} = 1,58$$

Deve ser observado que esta determinação do ganho também pode ser realizada facilmente sobre o gráfico log-módulo *versus* fase. A seguir demonstraremos como o diagrama log-módulo *versus* fase pode ser usado para determinar o ganho K de modo que o sistema apresente o valor desejado de M_r.

A Fig. 9.82 mostra o lugar geométrico de $M_r = 1,4$ e o lugar geométrico de $G(j\omega)/K$. A modificação do ganho não afeta o ângulo de fase, mas simplesmente movimenta a curva verticalmente para cima se $K > 1$ e para baixo se $K < 1$. Na Fig. 9.82, o lugar geométrico de $G(j\omega)/K$ deve ser deslocado para cima de 4 db para que seja tangente ao lugar geométrico M_r desejado e o lugar geométrico de $G(j\omega)/K$ inteiro possa estar fora do lugar geométrico $M_r = 1,4$. O deslocamento vertical, numericamente, do lugar geométrico de $G(j\omega)/K$ determina o ganho necessário para resultar no valor desejado de M_r. Portanto, resolvendo

$$20 \log K = 4$$

obtemos

$$K = 1,58$$

9.9 DETERMINAÇÃO EXPERIMENTAL DE FUNÇÕES DE TRANSFERÊNCIA

O primeiro passo na análise e projeto de um sistema de controle é deduzir um modelo matemático do processo sob consideração. A obtenção analítica de um modelo pode ser bem difícil. Devemos então obtê-lo por meio de análise experimental. A importância dos métodos de resposta em freqüência está no fato de que a função de transferência do processo, ou qualquer outro componente de um sistema, pode ser determinada através de medidas simples de resposta em freqüência.

Se forem medidas as relações de amplitude e a defasagem em um número suficiente de freqüências dentro da faixa de interesse, os valores podem ser utilizados para determinar o diagrama de Bode. A função de transferência pode então ser determinada através de aproximações assintóticas. Constroem-se as curvas assintóticas do log-módulo consistindo em vários segmentos e normalmente é possível, com algumas tentativas no processo de tentativa-e-erro, determinar as freqüências de canto e uma curva muito próxima à real. (Note que se a freqüência é indicada em ciclos por segundo e não em radianos por segundo, as freqüências de canto devem ser convertidas em radianos por segundo antes de se calcular as constantes de tempo.)

Geradores de sinais senoidais. Na realização de um teste de resposta em freqüência devem ser disponíveis geradores de sinais senoidais convenientes. O sinal pode estar na forma mecânica, elétrica ou pneumática. As faixas de freqüência necessárias para o teste são aproximadamente 0,001-10 cps para sistemas com grandes constantes de tempo e 0,1-1.000 cps para sistemas com pequenas constantes de tempo. O sinal senoidal deve ser razoavelmente isento de harmônicas ou distorção.

Para faixas de freqüência muito baixas (abaixo de 0,01 cps), um gerador de sinal mecânico (conjuntamente com um transdutor pneumático ou elétrico conveniente, se necessário) pode ser utilizado. Para a faixa de freqüência 0,01-1.000 cps, pode ser utilizado um gerador de sinal elétrico adequado (se necessário, juntamente com um transdutor conveniente).

Determinação de funções de transferência de mínima fase a partir de diagramas de Bode. Conforme estabelecido anteriormente, a verificação se um sistema é ou não de mínima fase, pode ser determinada a partir de curvas de resposta em freqüência examinando-se as características de alta freqüência.

Para determinar a função de transferência, inicialmente desenhamos as assíntotas à curva do log-módulo experimentalmente obtida. As assíntotas devem possuir inclinações múltiplas de $\pm$ 20 db/década. Se a inclinação da curva log-módulo obtida experimentalmente variar de -20 para -40 db/década em $\omega = \omega_1$, é claro que existe um fator $1/[1 + j(\omega/\omega_1)]$ na junção de transferência. Se a inclinação varia de -40 db/década em $\omega = \omega_2$, deve existir um fator quadrático da forma

$$\frac{1}{1 + 2\zeta\left(j\frac{\omega}{\omega_2}\right) + \left(j\frac{\omega}{\omega_2}\right)^2}$$

na função de transferência. A freqüência natural não amortecida deste fator quadrático é igual à freqüência de canto ω_2. A relação de amortecimento ζ pode ser determinada a partir da curva do log-módulo obtida experimentalmente medindo-se o valor do pico ressonante próximo à freqüência de canto ω_2 e comparando esta curva com aquela fornecida na Fig. 9.11.

Uma vez determinados os fatores da função de transferência $G(j\omega)$, o ganho pode ser determinado da parte de baixa-freqüência da curva do log-módulo. Como os termos do tipo $1 + j(\omega/\omega_1)$ e $1 + 2\zeta(j\omega/\omega_2) + (j\omega/\omega_2)^2$ tornam-se unitários conforme ω tende a zero, em freqüências muito baixas, a função de transferência senoidal $G(j\omega)$ pode ser escrita

$$\lim_{\omega \to 0} G(j\omega) = \frac{K}{(j\omega)^\lambda}$$

Na maioria dos sistemas práticos, λ é igual a 0, 1 ou 2.

1. Para $\lambda = 0$, ou sistemas tipo 0,

$$G(j\omega) = K \quad \text{para } \omega \ll 1$$

ou

$$20\,|\log G(j\omega)| = 20 \log K \quad \text{para } \omega \ll 1$$

A assíntota de baixa freqüência é uma reta horizontal em $20 \log K$ db. O valor de K pode então ser determinado a partir desta assíntota horizontal.

2. Para $\lambda = 1$, ou sistemas tipo 1,

$$G(j\omega) = \frac{K}{j\omega} \quad \text{para } \omega \ll 1$$

ou

$$20 \log |G(j\omega)| = 20 \log K - 20 \log \omega \quad \text{para } \omega \ll 1$$

o que indica que a assíntota de baixa-freqüência possui inclinação −20 db/década. A freqüência na qual a assíntota de baixa-freqüência (ou seu prolongamento) intercepta a reta 0-db é numericamente igual a K.

3. Para $\lambda = 2$, ou sistemas tipo 2.

$$G(j\omega) = \frac{K}{(j\omega)^2} \qquad \text{para } \omega \ll 1$$

ou

$$20 \log |G(j\omega)| = 20 \log K - 40 \log \omega \qquad \text{para } \omega \ll 1$$

A inclinação da assíntota em baixa-freqüência é −40 db/década. A freqüência na qual esta assíntota (ou seu prolongamento) intercepta a reta 0-db é numericamente igual a $\sqrt{K}$.

Exemplos de curvas log-módulo para sistemas tipo 0, tipo 1 e tipo 2 são fornecidos na Fig. 9.83, junto com a freqüência para a qual o ganho K está relacionado.

A curva do ângulo de fase obtida experimentalmente possibilita um meio de verificação da função de transferência obtida da curva do log-módulo. Para um sistema de mínima fase, a curva experimental do ângulo de fase deve concordar com precisão razoável com a curva teórica do ângulo de fase obtida a partir da função de transferência determinada. Estas duas curvas de ângulo de fase devem concordar exatamente tanto na faixa de freqüências muito altas como na faixa de freqüências muito baixas. Se o ângulo de fase obtido experimentalmente, em freqüências muito altas (comparadas com as freqüências de canto), não for igual a $-90°(q-p)$, onde p e q são os graus dos polinômios do numerador e do denominador da função de transferência, deve ser uma função de transferência de fase não-mínima.

Funções de transferência de fase não-mínima. Se, no extremo de alta freqüência, o ângulo de fase calculado for 180° menor do que o ângulo de fase obtido experimentalmente, então um dos zeros da função de transferência deve estar no semiplano direito do plano s, ao invés do semiplano esquerdo do plano s.

Se o ângulo de fase calculado diferir do ângulo de fase experimentalmente obtido por uma taxa de variação de fase constante, então existe um atraso de transporte, ou tempo morto, no sistema. Se supusermos que a função de transferência seja da forma

$$G(s)e^{-Ts}$$

onde $G(s)$ é a relação de dois polinômios em s, então

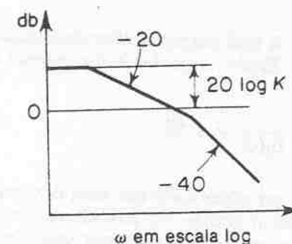
$$\lim_{\omega \to \infty} \frac{d}{d\omega} \underline{/G(j\omega)e^{-j\omega T}} = -T$$

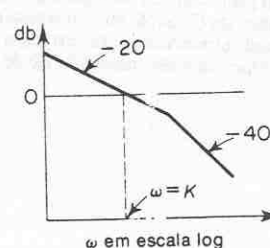
Portanto, desta última equação podemos calcular o valor do atraso de transporte T.

Algumas observações sobre a determinação experimental de funções de transferência

1. Normalmente é mais fácil fazer medidas precisas de amplitude do que de defasagem. Medidas de defasagem podem envolver erros causados por instrumentação ou por interpretação dos resultados experimentais.

2. A resposta em freqüência do equipamento de medida utilizado para medir





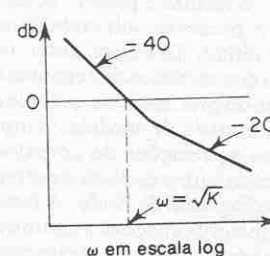


Fig. 9.83 (a) Curva log-módulo de um sistema tipo 0; (b) curvas log-módulo de sistemas tipo 1; (c) curvas log-módulo de sistemas tipo 2. (As inclinações indicadas estão em db/década.)

a saída do sistema deve possuir curvas de módulo *versus* freqüência praticamente planas. Além disso, o ângulo de fase deve ser aproximadamente proporcional à freqüência.

3. Sistemas físicos possuem alguns tipos de não linearidade. Portanto, é necessário considerar cuidadosamente a amplitude dos sinais de entrada senoidais. Se a amplitude do sinal de entrada for muito grande, o sistema saturará e o teste de resposta em freqüência apresentará resultados imprecisos. Por outro lado, um sinal pequeno causará erros devido à zona morta.

Conseqüentemente. deve ser feita uma escolha cuidadosa da amplitude do sinal de entrada senoidal. É necessário visualizar a forma de onda da saída do sistema para assegurar-se que a forma de onda é senoidal e que o sistema está operando na região linear durante todo o teste. (A forma de onda da saída do sistema não será senoidal se o sistema estiver operando em sua região não linear.)

4. Se o sistema sob consideração estiver operando continuamente por dias e semanas, então a operação normal não necessita ser interrompida para os testes de resposta em freqüência. O sinal de teste senoidal deve ser sobreposto às entradas normais. Então, para sistemas lineares, a saída devida ao sinal de teste estará sobreposta à saída normal. Para a determinação da função de transferência enquanto o sistema está em operação normal, também são muito utilizados sinais estocásticos (sinais de ruído branco). Utilizando funções de correlação, pode ser determinada a função de transferência do sistema sem interrupção na operação normal.

Exemplo 9.16 Determine a função de transferência do sistema cujas curvas de resposta em freqüência experimentais são indicadas na Fig. 9.84.

O primeiro passo na determinação da função de transferência é aproximar a curva do log-módulo por assíntotas com inclinações ± 20 db/década e seus múltiplos, conforme indicado na Fig. 9.84. As freqüências de canto são então determinadas e a seguinte forma da função de transferência é estabelecida:

$$G(j\omega) = \frac{K(1 + 0{,}5 j\omega)}{j\omega(1 + j\omega)\left[1 + 2\zeta\left(j\frac{\omega}{8}\right) + \left(j\frac{\omega}{8}\right)^2\right]}$$

O valor da relação de amortecimento ζ é determinado examinando-se o pico de ressonância próximo a $\omega = 6$ rad/s. Referindo-se à Fig. 9.11, ζ é determinada como sendo 0,5. O ganho K é numericamente igual à freqüência da intersecção do prolongamento da assíntota de baixa-freqüência com a reta 0-db. O valor de K é determinado em 10. Portanto, $G(j\omega)$ é determinada como uma tentativa para

$$G(j\omega) = \frac{10(1 + 0{,}5 j\omega)}{j\omega(1 + j\omega)\left[1 + \left(j\frac{\omega}{8}\right) + \left(j\frac{\omega}{8}\right)^2\right]}$$

ou

$$G(s) = \frac{320(s + 2)}{s(s + 1)(s^2 + 8s + 64)}$$

Esta função de transferência é uma tentativa porque não examinamos ainda a curva do ângulo de fase.

Uma vez verificadas as freqüências de canto da curva do log-módulo, a curva do ângulo de fase correspondente a cada fator da função de transferência pode ser desenhada facilmente. A soma destas curvas de ângulo de fase componentes é aquela da função de transferência suposta. A curva do ângulo de fase para $G(j\omega)$ é indicada por $\underline{/G}$ na Fig. 9.84. Desta figura notamos claramente uma discrepância entre a curva do ângulo de fase calculada e a curva do ângulo de fase obtida experimentalmente. A diferença entre as duas curvas em freqüências muito altas parece possuir uma taxa de variação constante. Conseqüentemente, as discrepâncias nas curvas de ângulo de fase devem ser causadas por um atraso de transporte.

Deste modo, suponhamos uma função de transferência completa como sendo $G(s)\,e^{-Ts}$.

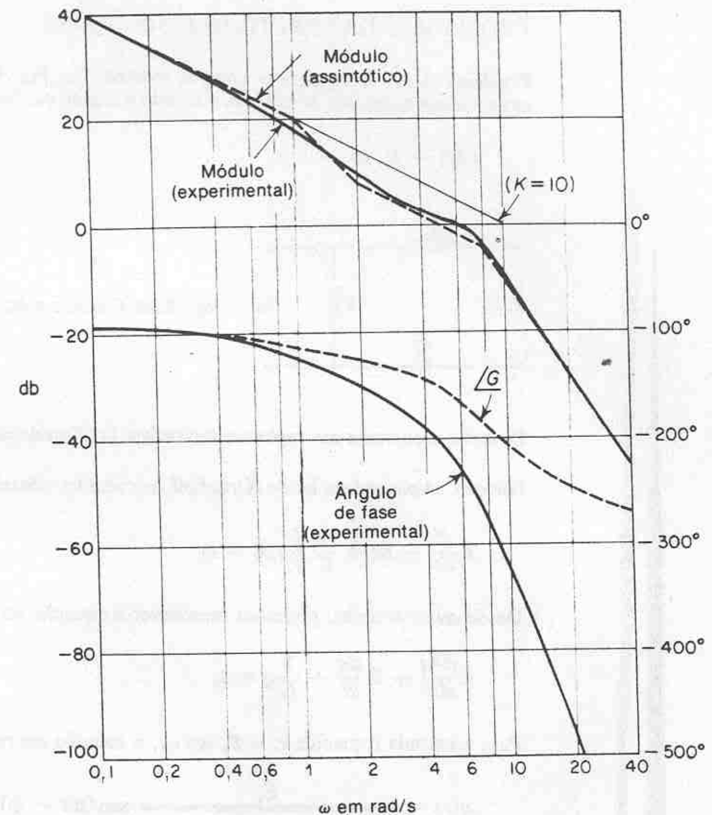


Fig. 9.84 Diagrama de Bode de um sistema. (Curvas cheias são curvas obtidas experimentalmente.)

Como a discrepância entre os ângulos de fase calculado e experimental é $-0{,}2\,\omega$ rad para freqüências muito altas, podemos determinar o valor de T como segue:

$$\lim_{\omega \to \infty} \frac{d}{d\omega} \underline{/G(j\omega)e^{-j\omega T}} = -T = -0{,}2$$

ou

$$T = 0{,}2 \text{ s}$$

A presença do atraso de transporte pode então ser determinada, e a função de transferência completa, determinada a partir das curvas experimentais, é

$$G(s)e^{-Ts} = \frac{320(s + 2)e^{-0{,}2s}}{s(s + 1)(s^2 + 8s + 64)}$$

PROBLEMAS ILUSTRATIVOS E SOLUÇÕES

Problema A.9.1 Considere o circuito indicado na Fig. 9.85. Suponha que a tensão e_i é aplicada aos terminais de entrada e resulta a tensão e_0. Suponha que a entrada do sistema é

$$e_i(t) = E_i \text{ sen } \omega t$$

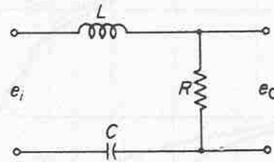


Fig. 9.85 Circuito elétrico.

Obtenha a corrente em regime estacionário $i(t)$ fluindo através da resistência R.

Solução. Aplicando a lei de Kirchhoff ao circuito, obtemos

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int i\,dt = e_i$$

Desde que $i = dq/dt$, podemos reescrever a equação do sistema como

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{1}{C}q = e_i$$

Para a entrada fornecida $e_i = E_i$ sen ωt, a solução em regime estacionário é

$$q(t) = \frac{E_i}{\sqrt{\left(\frac{1}{C} - L\omega^2\right)^2 + (R\omega)^2}} \text{ sen}(\omega t - \phi)$$

onde

$$\phi = \frac{R\omega}{\frac{1}{C} - L\omega^2}$$

Então a corrente $i(t)$ é

$$i(t) = \frac{dq(t)}{dt} = \frac{E_i\omega}{\sqrt{\left(\frac{1}{C} - L\omega^2\right)^2 + (R\omega)^2}} \cos(\omega t - \phi)$$

Problema A.9.2 Considere um sistema cuja função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{10(s+1)}{(s+2)(s+5)}$$

(Este sistema é idêntico ao considerado no Problema A.8.7.) É claro que os pólos de malha-fechada estão localizados em $s = -2$ e $s = -5$, e o sistema não é oscilatório. (A resposta

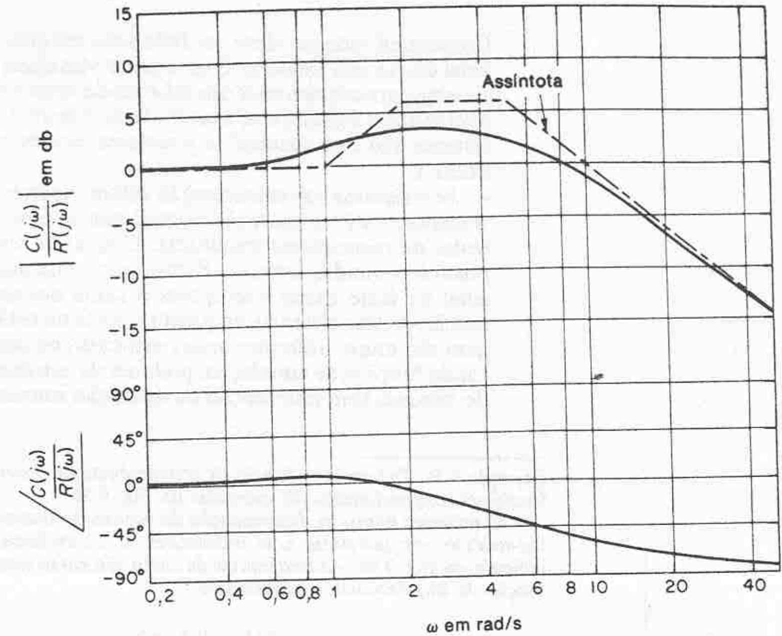


Fig. 9.86 Diagrama de Bode para $10(1 + j\omega)/[(2 + j\omega)(5 + j\omega)]$.

ao degrau unitário, entretanto, exibirá sobrelevação devido à presença de um zero em $s = -1$. Vide Fig. 8.48.)

Mostre que a resposta em freqüência de malha-fechada deste sistema exibirá um pico de ressonância, embora a relação de amortecimento dos pólos de malha-fechada seja maior do que a unidade.

Solução. A Fig. 9.86 mostra o diagrama de Bode para um sistema. O valor do pico de ressonância é aproximadamente 3,5 db. (Note que, na ausência de um zero, o sistema de segunda-ordem com $\zeta > 0,7$ não exibirá um pico de ressonância; entretanto, a presença de um zero de malha-fechada causará este pico.)

Problema A.9.3 Prove que o gráfico polar da função de transferência senoidal

$$G(j\omega) = \frac{j\omega T}{1 + j\omega T} \qquad (0 \leq \omega \leq \infty)$$

é uma semicircunferência. Determine o centro e o raio da circunferência.

Solução. A função de transferência senoidal dada $G(j\omega)$ pode ser escrita como segue:

$$G(j\omega) = X + jY$$

onde

$$X = \frac{\omega^2T^2}{1 + \omega^2T^2}, \qquad Y = \frac{\omega T}{1 + \omega^2T^2}$$

Então,

$$\left(X-\frac{1}{2}\right)^2+Y^2=\frac{(\omega^2T^2-1)^2}{4(1+\omega^2T^2)^2}+\frac{\omega^2T^2}{(1-\omega^2T^2)^2}=\frac{1}{4}$$

Portanto verificamos que o gráfico de $G(j\omega)$ é uma circunferência com centro em (0,5; 0) e raio igual a 0,5. A semicircunferência superior corresponde a $0 \leq \omega \leq \infty$, e a semicircunferência inferior corresponde a $-\infty \leq \omega \leq 0$.

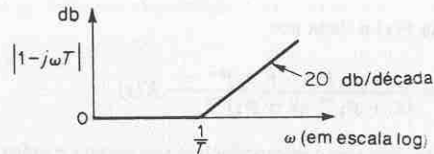
Problema A.9.4 Desenhe o diagrama de Bode do seguinte sistema de fase não-mínima:

$$\frac{C(s)}{R(s)}=1-Ts$$

Obtenha a resposta à rampa unitária do sistema e construa o gráfico de *c(t) versus t*.

Solução. O diagrama de Bode do sistema é indicado na Fig. 9.87.
Para uma entrada em rampa unitária, $R(s) = 1/s^2$, e

$$C(s)=\frac{1-Ts}{s^2}=\frac{1}{s^2}-\frac{T}{s}$$



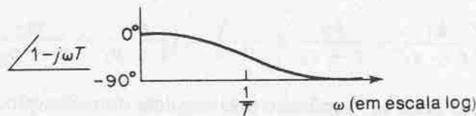


Fig. 9.87 Diagrama de Bode de $1 - j\omega T$.

A transformada de Laplace inversa de *C(s)* fornece

$$c(t) = t - T \qquad (t \geq 0)$$

A Fig. 9.88 mostra a curva de resposta *c(t) versus t*. (Note o comportamento defeituoso no início da resposta.) Uma propriedade característica de um sistema de fase não-mínima é que a

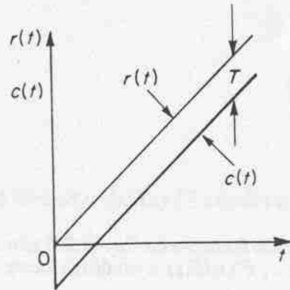
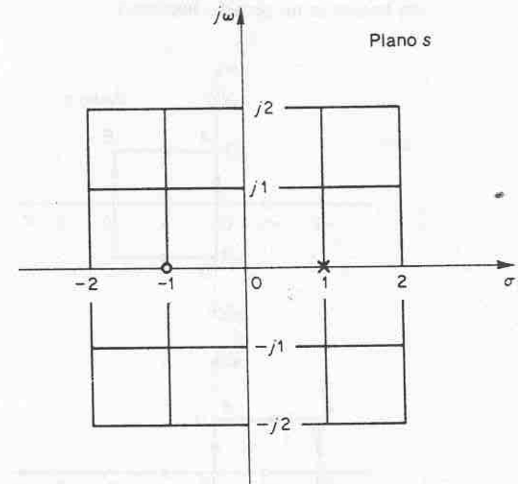


Fig. 9.88 Resposta à rampa unitária do sistema considerado no Problema A.9.4.

resposta transitória tem início no sentido oposto ao da entrada, porém, eventualmente retorna ao mesmo sentido.

Problema A.9.5 Considere a função

$$F(s)=\frac{s+1}{s-1}$$



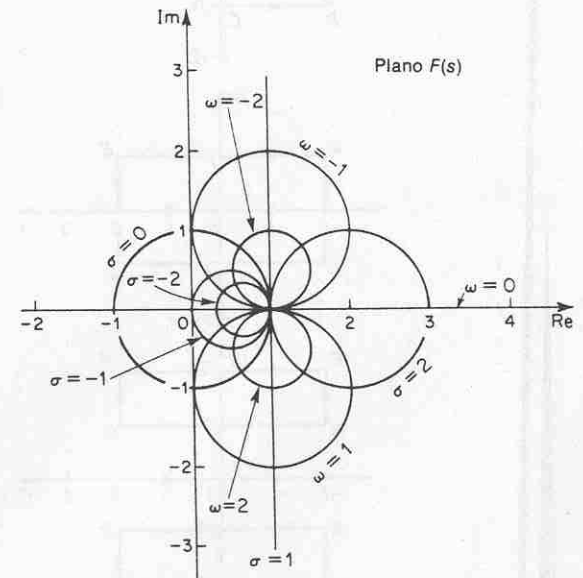


Fig. 9.89 Mapeamento conforme de grades de plano *s* no plano *F(s)* onde $F(s) = (s+1)/(s-1)$.

O mapeamento conforme das retas $\omega = 0, \pm 1, \pm 2$ e das retas $\sigma = 0, \pm 1, \pm 2$ fornece circunferências no plano $F(s)$, como indicado na Fig. 9.89. Mostre que se o contorno no plano s envolve o pólo de $F(s)$, há um envolvimento da origem no plano $F(s)$ no sentido anti-horário. Se o contorno no plano s envolve o zero de $F(s)$, há um envolvimento da origem do plano $F(s)$ no sentido horário. Se o contorno no plano s envolve tanto o zero como o pólo, ou se o contorno não envolve nem o zero nem o pólo, então não há envolvimento da origem do plano $F(s)$ pelo lugar geométrico de $F(s)$. (Note que no plano s, um ponto s representativo percorre um contorno no sentido horário.)

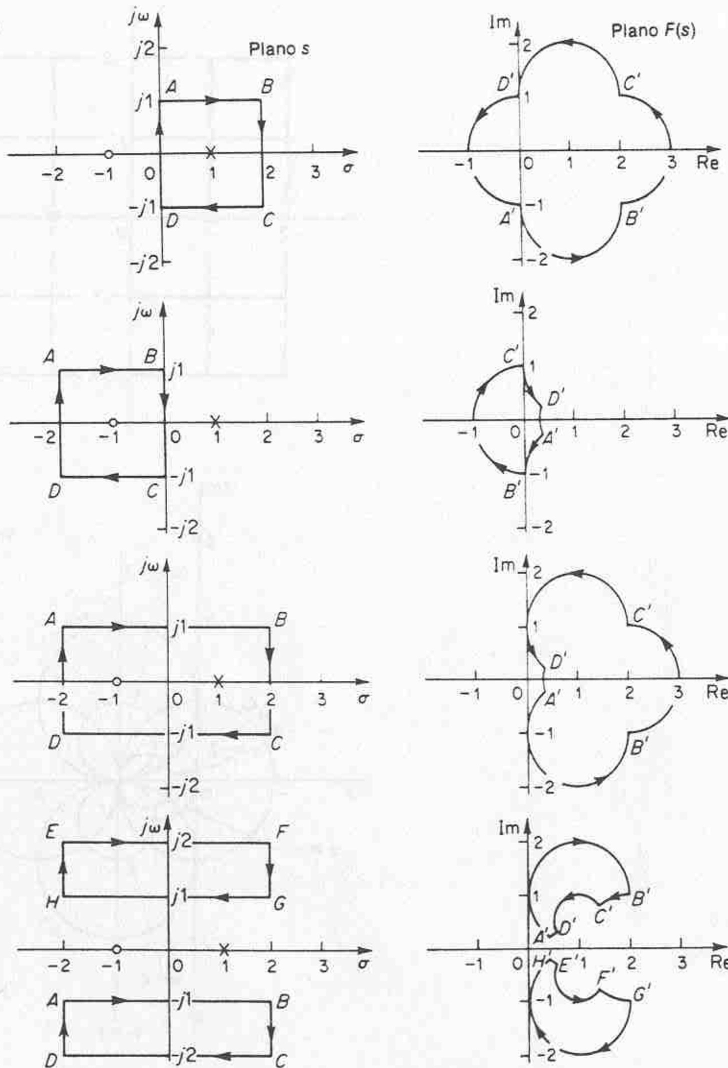


Fig. 9.90 Mapeamento conforme dos contornos no plano s para o plano $F(s)$ onde $F(s) = (s + 1)/(s - 1)$.

Solução. Uma solução gráfica é fornecida na Fig. 9.90; esta figura mostra contornos fechados no plano s e suas correspondentes curvas fechadas no plano $F(s)$.

Problema A.9.6 Prove o seguinte teorema de mapeamento: Seja $F(s)$ uma relação de polinômios em s. Seja P o número de pólos e Z o número de zeros de $F(s)$ que existem dentro de um contorno fechado no plano s, levando-se em conta a multiplicidade. Seja o contorno fechado tal que não passa através de quaisquer pólos ou zeros de $F(s)$. O contorno fechado no plano s mapeia então o plano $F(s)$ com uma curva fechada. O número N de envolvimentos no sentido horário da origem do plano $F(s)$, conforme o ponto representativo s percorre o contorno inteiro no plano s no sentido horário, é igual a $Z - P$.

Solução. Para provar este teorema, usaremos o teorema de Cauchy e o teorema do resíduo. O teorema de Cauchy estabelece que a integral de $F(s)$ ao longo de um contorno fechado no plano s é nula se $F(s)$ é analítica dentro e sobre o contorno fechado, ou

$$\oint F(s)\,ds = 0$$

O teorema do resíduo estabelece que a integral de $F(s)$ tomada no sentido horário sobre o contorno fechado no plano s é igual a $-2\pi j$ vezes os resíduos nos pólos simples de $F(s)$, ou

$$\oint F(s)\,ds = -2\pi j\,(\sum \text{resíduos})$$

Suponha que $F(s)$ é dada por

$$F(s) = \frac{(s + z_1)^{k_1}(s + z_2)^{k_2}\cdots}{(s + p_1)^{m_1}(s + p_2)^{m_2}\cdots}X(s)$$

onde $X(s)$ é analítica no contorno fechado do plano s e todos os pólos e zeros estão localizados no contorno. Então, a relação $F'(s)/F(s)$ pode ser escrita

$$\frac{F'(s)}{F(s)} = \left(\frac{k_1}{s + z_1} + \frac{k_2}{s + z_2} + \cdots\right) - \left(\frac{m_1}{s + p_1} + \frac{m_2}{s + p_2} + \cdots\right) + \frac{X'(s)}{X(s)} \quad (9\text{-}24)$$

Este resultado pode ser verificado pela seguinte consideração: Se $F(s)$ é dado por

$$F(s) = (s + z_1)^k X(s)$$

então $F(s)$ possui um zero de k-ésima ordem em $s = -z_1$. Diferenciando $F(s)$ em relação a s, resulta

$$F'(s) = k(s + z_1)^{k-1}X(s) + (s + z_1)^k X'(s)$$

Portanto,

$$\frac{F'(s)}{F(s)} = \frac{k}{s + z_1} + \frac{X'(s)}{X(s)} \quad (9\text{-}25)$$

Verificamos que considerando a relação $F'(s)/F(s)$, o zero de k-ésima ordem de $F(s)$ resulta um pólo simples de $F'(s)/F(s)$.

Se o último termo do segundo membro da Eq. (9.25) não contém qualquer pólo ou zero no contorno fechado do plano s, $F'(s)/F(s)$ é analítica neste contorno exceto no zero em $s = -z_1$. Então,

$$\oint \frac{F'(s)}{F(s)}\,ds = \oint \frac{k}{s+z_1} + \oint \frac{X'(s)}{X(s)} = -2\pi jk$$

Referindo-se à Eq. (9.24) e notando que $X'(s)/X(s)$ é analítica no contorno e que os fatores nos parênteses são todos pólos simples localizados no contorno, obtemos a seguinte relação:

$$\oint \frac{F'(s)}{F(s)}\,ds = -2\pi j[(k_1 + k_2 + \cdots) - (m_1 + m_2 + \cdots)] = -2\pi j(Z - P)$$

onde

$Z = k_1 + k_2 + \ldots$ = número total de zeros de $F(s)$ envolvidos no contorno fechado do plano s
$P = m_1 + m_2 + \ldots$ = número total de pólos de $F(s)$ envolvidos no contorno fechado do plano s

(Os zeros ou pólos com multiplicidade k são considerados k zeros ou pólos localizados no mesmo ponto.) Desde que $F(s)$ é um número complexo, $F(s)$ pode ser escrito

$$F(s) = |F|e^{j\theta}$$

e

$$\ln F(s) = \ln|F| + j\theta$$

Notando que $F'(s)/F(s)$ pode ser escrita

$$\frac{F'(s)}{F(s)} = \frac{d\ln F(s)}{ds}$$

obtemos

$$\frac{F'(s)}{F(s)} = \frac{d\ln|F|}{ds} + j\frac{d\theta}{ds}$$

Se o contorno fechado no plano s é mapeado no contorno fechado Γ no plano $F(s)$, então

$$\oint \frac{F'(s)}{F(s)}\,ds = \oint_\Gamma d\ln|F| + j\oint_\Gamma d\theta = j\int d\theta = 2\pi j(P - Z)$$

A integral $\int_\Gamma d\ \ln|F|$ é nula desde que o valor de $\ln|F|$ é o mesmo no ponto inicial e no ponto final do contorno Γ. Portanto, obtemos

$$\frac{\theta_2 - \theta_1}{2\pi} = P - Z$$

A diferença angular entre os valores final e inicial de θ é igual à variação total no ângulo de fase de $F'(s)/F(s)$ conforme um ponto representativo no plano s se move ao longo do contorno fechado. Notando que N é o número de envolvimentos da origem do plano $F(s)$ no sentido horário e $\theta_2 - \theta_1$ é nulo ou um múltiplo de 2π rad, obtemos

$$\frac{\theta_2 - \theta_1}{2\pi} = -N$$

Portanto, temos a relação

$$N = Z - P$$

Isto prova o teorema.

Note que por meio deste teorema de mapeamento, os números exatos de zeros e de pólos não podem ser determinados, mas apenas a diferença entre eles. Note também que das Figs. 9.91(a) e (b) verificamos que se θ não varia através de 2π rad, então a origem do plano $F(s)$ não pode estar envolvida.

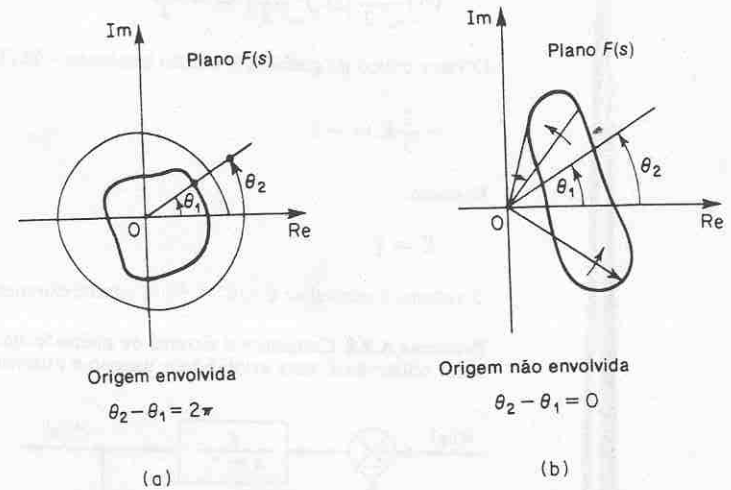


Fig. 9.91 Determinação do envolvimento da origem do plano $F(s)$.

Problema A.9.7 O sistema com a função de transferência de malha-aberta seguinte e com $K = 2$ é estável?

$$G(s)H(s) = \frac{K}{s(s+1)(2s+1)}$$

Determine o valor crítico do ganho K para estabilidade.

Solução. A função de transferência de malha-aberta é

$$G(j\omega)H(j\omega) = \frac{K}{j\omega(j\omega+1)(2j\omega+1)} = \frac{K}{-3\omega^2 + j\omega(1-2\omega^2)}$$

Esta função de transferência de malha-aberta não possui pólos no semiplano direito do plano s. Portanto, para estabilidade, o ponto $-1 + j0$ não deve estar envolvido pelo gráfico de Nyquist. Vamos determinar o ponto onde o gráfico de Nyquist cruza o eixo real negativo. Consideremos a parte imaginária de $G(j\omega)\,H(j\omega)$ como sendo nula, ou

$$1 - 2\omega^2 = 0$$

de onde obtemos

$$\omega = \pm\frac{1}{\sqrt{2}}$$

Substituindo $\omega = 1/\sqrt{2}$ em $G(j\omega)\,H(j\omega)$, obtemos

$$G\left(j\frac{1}{\sqrt{2}}\right)H\left(j\frac{1}{\sqrt{2}}\right) = -\frac{2K}{3}$$

O valor crítico do ganho K é obtido igualando $-2K/3$ a -1, ou

$$-\frac{2}{3}K = -1$$

Portanto,

$$K = \tfrac{3}{2}$$

O sistema é estável se $0 < K < 3/2$. Conseqüentemente, o sistema com $K = 2$ é instável.

Problema A.9.8 Considere o sistema de malha-fechada indicado na Fig. 9.92. Determine o valor crítico de K para estabilidade usando o critério de Nyquist.

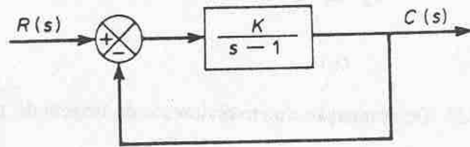


Fig. 9.92 Sistema de malha-fechada.

Solução. O gráfico polar de

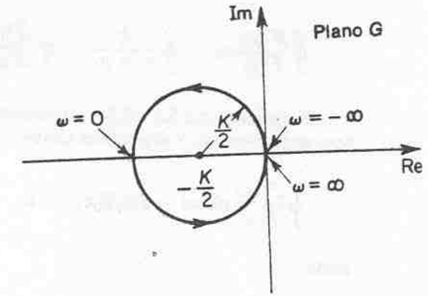
$$G(j\omega) = \frac{K}{j\omega - 1}$$

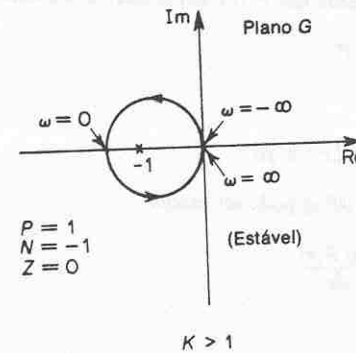
é uma circunferência com centro em $-K/2$ sobre o eixo real negativo e raio $K/2$, conforme indicada na Fig. 9.93(a). Conforme ω aumenta desde $-\infty$ até ∞, o lugar geométrico de $G(j\omega)$ faz uma rotação no sentido anti-horário. Neste sistema, $P = 1$ porque há um pólo de $G(s)$ no semiplano direito do plano s. Para o sistema de malha-fechada ser estável, Z deve ser igual a zero. Portanto, $N = Z - P$ deve ser igual a -1, ou deve haver um envolvimento no sentido anti-horário do ponto $-1 + j0$ para estabilidade. (Se não houver envolvimento do ponto $-1 + j0$, o sistema será instável.) Portanto, para estabilidade, K deve ser maior do que a unidade, e $K = 1$ fornece o limite de estabilidade. A Fig. 9.93(b) mostra os casos estável e instável dos gráficos de $G(j\omega)$.

Problema A.9.9 Considere o sistema com realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{Ke^{-0,8s}}{s+1}$$

Utilizando o gráfico de Nyquist, determine o valor crítico de K para estabilidade.





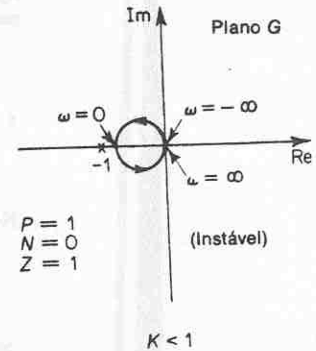


Fig. 9.93 (a) Gráfico polar de $K/(j\omega - 1)$; (b) gráficos polares de $K/(j\omega - 1)$ para os casos estável e instável.

Solução. Para este sistema,

$$G(j\omega) = \frac{Ke^{-0,8j\omega}}{j\omega + 1}$$

$$= \frac{K(\cos 0,8\omega - j\,\mathrm{sen}\,0,8\omega)(1 - j\omega)}{1 + \omega^2}$$

$$= \frac{K}{1+\omega^2}\left[(\cos 0,8\omega - \omega\,\mathrm{sen}\,0,8\omega) - j(\mathrm{sen}\,0,8\omega + \omega\cos 0,8\omega)\right]$$

A parte imaginária de $G(j\omega)$ é igual a zero se

$$\mathrm{sen}\,0,8\omega + \omega\cos 0,8\omega = 0$$

Portanto,

$$\omega = -\tan 0,8\omega$$

Resolvendo esta equação para o valor mínimo positivo de ω, obtemos

$$\omega = 2,45$$

Substituindo $\omega = 2,45$ em $G(j\omega)$, obtemos

$$G(j2,45) = \frac{K}{1 + 2,45^2}(\cos 1,96 - 2,45 \operatorname{sen} 1,96)$$
$$\doteq -0,378K$$

O valor crítico de K para estabilidade é obtido impondo-se que $G(j2,45)$ seja igual a -1. Portanto,

$$0,378K = 1$$

ou

$$K = 2,65$$

A Fig. 9.94 mostra os gráficos polares ou de Nyquist de $2,65\, e^{-0,8j\omega}/(1 + j\omega)$ e $2,65/(1 + j\omega)$. O sistema de primeira-ordem sem atraso de transporte é estável para todos os valores de K, porém o outro com atraso de transporte de 0,8 s resulta instável para $K > 2,65$.

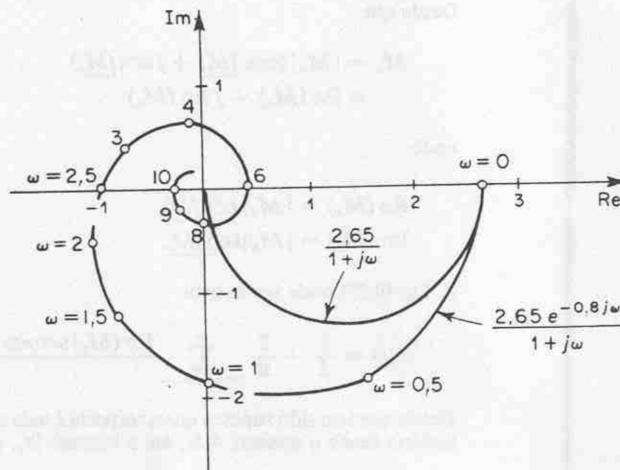


Fig. 9.94 Gráficos polares de $2,65e^{-0,8j\omega}/(1 + j\omega)$ e $2,65/(1 + j\omega)$.

Problema A.9.10* Consideraremos aqui um método para obtenção da resposta ao degrau diretamente dos dados de resposta em freqüência de malha-fechada. Neste método, uma entrada em degrau unitário é substituída por uma entrada em onda quadrada e a saída para a entrada com onda quadrada é aproximada por meio de uma série convergente de senos. As hipóteses consideradas neste método são sumariadas a seguir:

*Referência S-6.

1. O sistema é oscilatório. Em outras palavras, o método aplica-se somente a sistemas oscilatórios.
2. O instante do primeiro pico na resposta ao degrau é $t_1 = \pi/\omega_0$, onde ω_0 é a freqüência na qual $\underline{/C(j\omega)/R(j\omega)} = -90°$.
3. O sistema apresenta um pequeno erro desprezível para uma entrada em degrau unitário em $t \geq 4,5t_1$.

Para um sistema de segunda-ordem da forma

$$\frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

o erro em uma resposta ao degrau unitário é menor do que 2% para $t > 4,5t_1$ se a relação de amortecimento ζ for maior do que 0,267. Pode ser verificado que se o semiperíodo da onda quadrada for suficientemente longo, comparado com as constantes de tempo de sistema, então a substituição da entrada em degrau por uma entrada com onda quadrada não introduz restrições adicionais ao sistema. Para uma boa precisão, é muito importante uma escolha apropriada da freqüência da onda quadrada. No presente método, este período é escolhido como $18t_1$. A entrada com onda quadrada e a resposta do sistema correspondente são esboçadas na Fig. 9.95. O motivo desta escolha do período da onda quadrada é o seguinte: A resposta ideal de um sistema para uma entrada em degrau unitário deve ser aquela mostrada na Fig. 9. 96(a), onde t_1 é o intervalo de tempo para a resposta atingir o valor máximo pela primeira vez. A resposta ideal ao impulso unitário correspondente é indicada na Fig. 9.96(b). A transformada de Laplace da resposta impulsiva ideal é então

$$C(s) = \frac{1}{t_1 s}(1 - e^{-t_1 s})$$

Desde que a transformada de Laplace da entrada impulsiva unitária é igual a um, $R(s) = 1$, e

$$\frac{C(s)}{R(s)} = \frac{1}{t_1 s}(1 - e^{-t_1 s})$$

do qual

$$\frac{C(j\omega)}{R(j\omega)} = \frac{1}{j\omega t_1}(1 - e^{-j\omega t_1})$$
$$= \frac{\sqrt{2(1 - \cos \omega t_1)}}{\omega t_1} e^{j\theta}$$

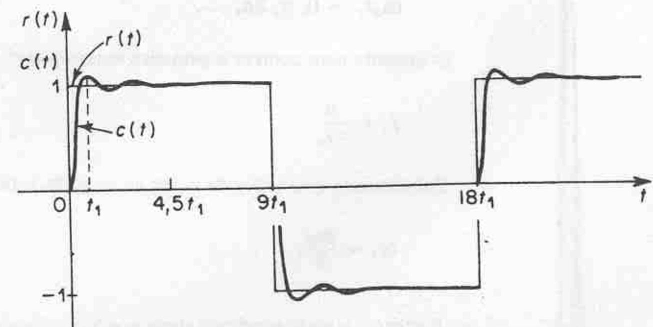


Fig. 9.95 Entrada com onda quadrada e a resposta correspondente de um sistema de segunda-ordem.

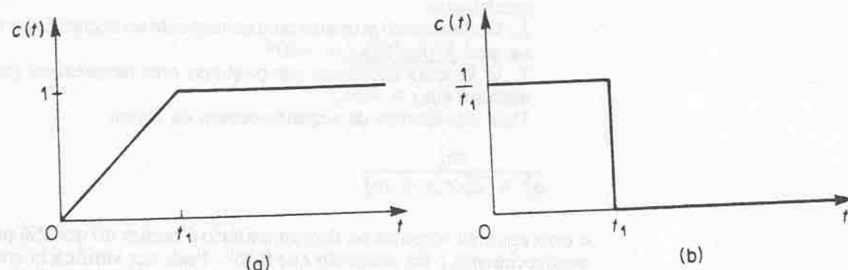


Fig. 9.96 (a) Resposta ideal ao degrau unitário; (b) resposta ideal ao impulso unitário.

onde

$$\theta = -\frac{\pi}{2} + \text{tg}^{-1}\frac{\text{sen}\,\omega t_1}{1 - \cos\omega t_1}$$

Em $\omega = \omega_0$, tem-se admitido que

$$\underline{/\frac{C(j\omega_0)}{R(j\omega_0)}} = -\frac{\pi}{2}$$

Portanto, em $\omega = \omega_0$,

$$\theta = -\frac{\pi}{2} = -\frac{\pi}{2} + \text{tg}^{-1}\frac{\text{sen}\,\omega_0 t_1}{1 - \cos\omega_0 t_1}$$

de onde

$$\text{sen}\,\omega_0 t_1 = 0$$

ou

$$\omega_0 t_1 = 0, \pi, 2\pi, \ldots$$

O instante para ocorrer o primeiro máximo tem sido admitido como

$$t_1 = \frac{\pi}{\omega_0}$$

Substituindo este valor de t_1 em $\omega_1 = \pi/(9t_1)$, obtemos

$$\omega_1 = \frac{\omega_0}{9} \tag{9-26}$$

Portanto, a onda quadrada deve possuir a freqüência $\omega_0/9$ (ou o período $18t_1$), onde ω_0 é a freqüência na qual a resposta em freqüência de malha-fechada exibe um atraso de fase de 90°. Se a freqüência da onda quadrada for muito alta, a precisão não será boa, enquanto uma freqüência demasiadamente baixa resulta em uma convergência lenta da série.

A entrada com a onda quadrada $r(t)$ mostrada na Fig. 9.95 pode ser expandida na seguinte série de Fourier:

$$r(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1,3,5,\ldots}^{\infty}\frac{\text{sen}\,n\omega_1 t}{n}$$

onde

$$\omega_1 = \frac{\pi}{9t_1}$$

A resposta do sistema para esta entrada com uma onda quadrada é a soma das respostas para cada componente da série de Fourier. Se considerarmos

$$\frac{C(j\omega)}{R(j\omega)} = M(j\omega)$$

então,

$$c(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1,3,5,\ldots}^{\infty}|M_n|\frac{\text{sen}\,(n\omega_1 t + \underline{/M_n})}{n} \tag{9-27}$$

onde

$$M_n = M(jn\omega_1)$$

Desde que

$$\begin{aligned} M_n &= |M_n|(\cos\underline{/M_n} + j\,\text{sen}\,\underline{/M_n}) \\ &= \text{Re}\,(M_n) + j\,\text{Im}\,(M_n) \end{aligned}$$

onde

$$\begin{aligned} \text{Re}\,(M_n) &= |M_n|\cos\underline{/M_n} \\ \text{Im}\,(M_n) &= |M_n|\,\text{sen}\,\underline{/M_n} \end{aligned}$$

A Eq. (9.27) pode ser escrita

$$c(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1,3,5,\ldots}^{\infty}\frac{\text{Re}\,(M_n)\,\text{sen}\,n\omega_1 t + \text{Im}\,(M_n)\cos n\omega_1 t}{n} \tag{9-28}$$

Desde que tem sido suposto que a resposta é nula desde o instante $-4{,}5t_1$ até o instante zero e unitária desde o instante $4{,}5t_1$ até o instante $9t_1$, substituindo $-t$ na Eq. (9.28), resulta

$$c(-t) = 0 = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1,3,5,\ldots}^{\infty}\left[\frac{-\text{Re}\,(M_n)\,\text{sen}\,n\omega_1 t}{n} + \frac{\text{Im}\,(M_n)\cos n\omega_1 t}{n}\right] \tag{9-29}$$
$$(0 \le t \le 4{,}5t_1)$$

Subtraindo a Eq. (9.29) da Eq. (9.28), resulta

$$c(t) = \frac{4}{\pi}\sum_{n=1,3,5,\ldots}^{\infty}\frac{\text{Re}\,(M_n)\,\text{sen}\,n\omega_1 t}{n} \qquad (0 \le t \le 4{,}5t_1) \tag{9-30}$$

Da Eq. (9.30), podemos calcular a resposta ao degrau unitário diretamente dos dados da resposta em freqüência de malha-fechada.

Na prática, o processo de soma deve ser truncado. Geralmente, os primeiros sete ou oito termos da série infinita na Eq. (9.30) serão suficientes para assegurar uma precisão adequada. A convergência e a precisão podem ser testadas facilmente colocando-se $\omega_1 t$ igual a 90º e calculando a resposta como segue:

$$c\left(\frac{\pi}{2\omega_1}\right) = \frac{4}{\pi}\Big[\operatorname{Re}(M_1) - \frac{1}{3}\operatorname{Re}(M_3) + \frac{1}{5}\operatorname{Re}(M_5) - \frac{1}{7}\operatorname{Re}(M_7) - \frac{1}{11}\operatorname{Re}(M_{11}) + \frac{1}{13}\operatorname{Re}(M_{13}) - \frac{1}{15}\operatorname{Re}(M_{15}) + \cdots\Big] \tag{9-31}$$

A soma desta série deve ser igual à unidade. Na Eq. (9.31) o termo correspondente à nona harmônica de freqüência anula-se porque nesta freqüência o ângulo de fase de $C(j\omega)/R(j\omega)$ é $-90°$, de modo que a parte real de $C(j\omega)/R(j\omega)$ é nula. Os termos correspondentes às freqüências maiores são bastante pequenos e a série converge rapidamente.

Considere agora o lugar geométrico da resposta em freqüência de malha-fechada indicada na Fig. 9.97. Obtenha a máxima sobrelevação à resposta ao degrau unitário do sistema.

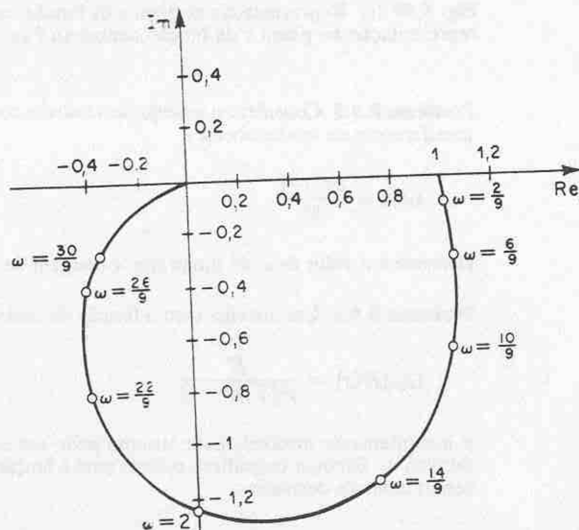


Fig. 9.97 Gráfico polar de uma resposta em freqüência de malha-fechada.

Solução. Do lugar geométrico fornecido, ω_0 é determinado como sendo 2. Da Eq. (9.26), ω_1 é

$$\omega_1 = \frac{2}{9}$$

Da Eq. (9.30), pode ser calculada a resposta ao degrau unitário e é igual a

$$c(t) \doteq \frac{4}{\pi}\Big[\operatorname{Re}(M_1)\operatorname{sen}\frac{2}{9}t + \frac{\operatorname{Re}(M_3)}{3}\operatorname{sen}\frac{6}{9}t + \frac{\operatorname{Re}(M_5)}{5}\operatorname{sen}\frac{10}{9}t + \frac{\operatorname{Re}(M_7)}{7}\operatorname{sen}\frac{14}{9}t + \frac{\operatorname{Re}(M_{11})}{11}\operatorname{sen}\frac{22}{9}t + \frac{\operatorname{Re}(M_{13})}{13}\operatorname{sen}\frac{26}{9}t + \frac{\operatorname{Re}(M_{15})}{15}\operatorname{sen}\frac{30}{9}t\Big] \tag{9-32}$$

Da Fig. 9.97, Re (M_1) pode ser determinado como segue:

$$\begin{aligned}
\operatorname{Re}(M_1) &= 1{,}02\\
\operatorname{Re}(M_3) &= 1{,}04\\
\operatorname{Re}(M_5) &= 1{,}03\\
\operatorname{Re}(M_7) &= 0{,}72\\
\operatorname{Re}(M_{11}) &= -0{,}40\\
\operatorname{Re}(M_{13}) &= -0{,}42\\
\operatorname{Re}(M_{15}) &= -0{,}36
\end{aligned}$$

A precisão da solução pode ser verificada substituindo-se estes valores de Re(M_i) e $t = (\pi/2)(9/2)$ na Eq. (9.32). A soma pode ser obtida como segue:

$$c\left(\frac{\pi}{2}\frac{9}{2}\right) = \frac{4}{\pi}\Big[1{,}02 - \frac{1{,}04}{3} + \frac{1{,}03}{5} - \frac{0{,}72}{7} - \frac{(-0{,}40)}{11} + \frac{(-0{,}42)}{13} - \frac{(-0{,}36)}{15}\Big] = 1{,}023$$

Este valor difere da unidade, porém esta discrepância está dentro dos erros de cálculo de acordo com as várias aproximações. A sobrelevação máxima na resposta ao degrau unitário ocorre em

$$t = \frac{\pi}{\omega_0} = \frac{\pi}{2}$$

Portanto,

$$c\left(\frac{\pi}{2}\right) = \frac{4}{\pi}\Big[1{,}02\operatorname{sen}\frac{\pi}{9} + \frac{1{,}04}{3}\operatorname{sen}\frac{\pi}{3} + \frac{1{,}03}{5}\operatorname{sen}\frac{5\pi}{9} + \frac{0{,}72}{7}\operatorname{sen}\frac{7\pi}{9} - \frac{0{,}40}{11}\operatorname{sen}\frac{11\pi}{9} - \frac{0{,}42}{13}\operatorname{sen}\frac{13\pi}{9} - \frac{0{,}36}{15}\operatorname{sen}\frac{15\pi}{9}\Big] = 1{,}27$$

$$\text{Máxima sobrelevação} = 1{,}27 - 1 = 0{,}27$$

A sobrelevação máxima é determinada em 27%.

O lugar geométrico indicado na Fig. 9.97 corresponde realmente àquele do seguinte sistema:

$$\frac{C(j\omega)}{R(j\omega)} = M(j\omega) = \frac{4}{(j\omega)^2 + 1{,}6(j\omega) + 4}$$

(Esta informação é aqui fornecida para verificação da precisão do resultado e não é necessária para a solução deste problema.) O valor exato da sobrelevação máxima para este sistema é 25,4%. A precisão da solução deste exemplo é normalmente adequada para fins de engenharia. Se for desejada a curva completa da resposta ao degrau unitário, ela pode ser construída a partir da Eq. (9.32).

Problema A.9.11 Um sistema de controle de malha-fechada pode incluir um elemento instável dentro da malha. Quando o critério de estabilidade de Nyquist é aplicado a um sistema deste tipo, devem ser obtidas as curvas de resposta em freqüência para o elemento instável.

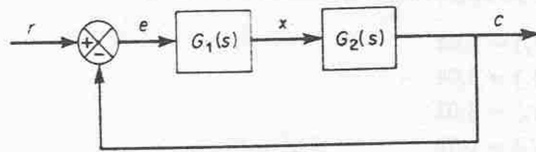


Fig. 9.98 Sistema de controle.

Como podemos obter experimentalmente as curvas de resposta em freqüência de um elemento instável? Sugira uma abordagem possível para a determinação experimental da resposta em freqüência de um elemento linear instável.

Solução. Uma abordagem possível é medir as características de resposta em freqüência de um elemento instável usando-o como uma parte de um sistema estável.

Considere o sistema indicado na Fig. 9.98. Suponha que o elemento $G_1(s)$ é instável. O sistema completo pode ser feito estável pela escolha de um conveniente elemento linear $G_2(s)$. Aplicamos um sinal senoidal na entrada. Em regime estacionário, todos os sinais na malha serão senoidais. Medimos os sinais $e(t)$, a entrada para o elemento instável, e $x(t)$, a saída do elemento instável. Modificando a freqüência [e possivelmente a amplitude para uma medida conveniente de $e(t)$ e $x(t)$] do sinal de entrada e repetindo este processo , é possível obter a resposta em freqüência do elemento instável.

PROBLEMAS

Problema B.9.1 Considere o sistema com realimentação unitária e função de transferência de malha-aberta

$$G(s) = \frac{10}{s+1}$$

Obter a saída em regime estacionário do sistema quando estiver sujeito a cada uma das seguintes entradas:

1. $r(t) = \text{sen}\,(t + 30°)$
2. $r(t) = 2\cos(2t - 45°)$
3. $r(t) = \text{sen}\,(t + 30°) - 2\cos(2t - 45°)$

Problema B.9.2 Considere o sistema cuja função de transferência de malha-fechada é

$$\frac{C(s)}{R(s)} = \frac{K(T_2 s + 1)}{T_1 s + 1}$$

Obter a saída em regime estacionário do sistema quando estiver sujeito à entrada $r(t) = R\,\text{sen}\,\omega t$.

Problema B.9.3 **Esboce os diagramas de Bode para as três funções de transferência seguintes:**

1. $G(s) = \dfrac{T_1 s + 1}{T_2 s + 1} \quad (T_1 > T_2 > 0)$
2. $G(s) = \dfrac{T_1 s - 1}{T_2 s + 1} \quad (T_1 > T_2 > 0)$
3. $G(s) = \dfrac{-T_1 s + 1}{T_2 s + 1} \quad (T_1 > T_2 > 0)$

Problema B.9.4 As configurações de pólos e zeros das funções complexas $F_1(s)$ e $F_2(s)$ são indicadas nas Figs. 9.99(a) e (b), respectivamente. Suponha que os contornos fechados no plano s são aqueles indicados nas Figs. 9.99(a) e (b). Esboçar qualitativamente os contornos fechados correspondentes no plano $F_1(s)$ e no plano $F_2(s)$.

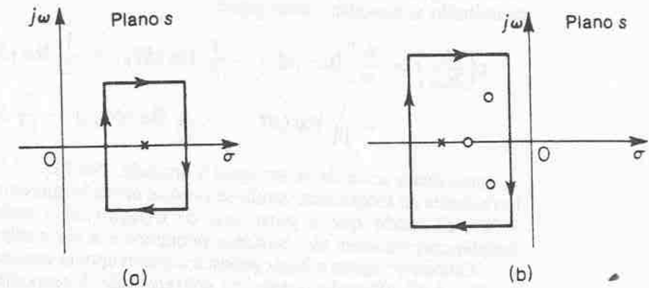


Fig. 9.99 (a) Representação no plano s da função complexa $F_1(s)$ e um contorno fechado; (b) representação no plano s da função complexa $F_2(s)$ e um contorno fechado.

Problema B.9.5 Considere o sistema de controle com realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{as+1}{s^2}$$

Determine o valor de a de modo que a margem de fase seja igual a 45°.

Problema B.9.6 Um sistema com a função de transferência de malha-aberta

$$G(s)H(s) = \frac{K}{s^2(T_1 s + 1)}$$

é inerentemente instável. Este sistema pode ser estabilizado adicionando-se um controle derivativo. Esboçar os gráficos polares para a função de transferência de malha-aberta com e sem o controle derivativo.

Problema B.9.7 Esboçar os gráficos polares da função de transferência de malha-aberta

$$G(s)H(s) = \frac{K(T_a s + 1)(T_b s + 1)}{s^2(T_1 s + 1)}$$

para os dois casos seguintes:

1. $T_a > T_1 > 0, \quad T_b > T_1 > 0$
2. $T_1 > T_a > 0, \quad T_1 > T_b > 0$

Problema B.9.8 Considere o sistema com

$$G(s) = \frac{10}{s(s-1)}, \qquad H(s) = 1 + K_h s$$

Determinar o valor crítico de K_h para a estabilidade do sistema de malha-fechada.

Problema B.9.9 Considere o sistema de malha-fechada com a seguinte função de transferência de malha-aberta:

$$G(s)H(s) = \frac{10K(s + 0{,}5)}{s^2(s + 2)(s + 10)}$$

Construir os gráficos polares direto e inverso de $G(s)H(s)$ com $K = 1$ e $K = 10$. Aplicar o critério de estabilidade de Nyquist para os gráficos e determinar a estabilidade do sistema para estes valores de K.

Problema B.9.10 Considere o sistema de malha-fechada cuja função de transferência de malha-aberta é

$$G(s)H(s) = \frac{Ke^{-2s}}{s}$$

Determinar o valor máximo de K para o qual o sistema é estável.

Problema B.9.11 Considere o sistema de malha-fechada cuja função de transferência de malha-aberta é

$$G(s)H(s) = \frac{Ke^{-Ts}}{s(s + 1)}$$

Determinar o valor máximo do ganho K para estabilidade em função do tempo morto T.

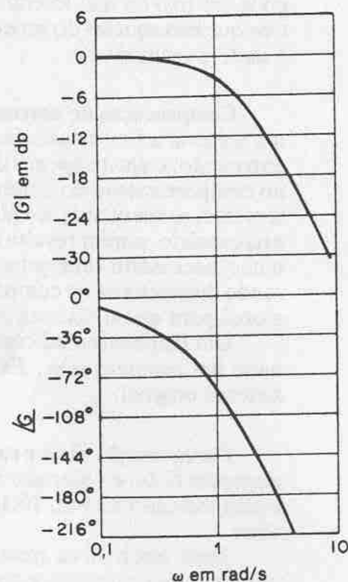


Fig. 9.100 Diagrama de Bode de um sistema determinado experimentalmente.

Problema B.9.12 Esboçar o gráfico polar de

$$G(s) = \frac{(Ts)^2 - 6(Ts) + 12}{(Ts)^2 + 6(Ts) + 12} \tag{9-33}$$

Mostrar que para a faixa de freqüência $0 < \omega T < 2\sqrt{3}$, a Eq. (9.33) fornece uma boa aproximação para o atraso de transporte e^{-Ts}.

Problema B.9.13 O diagrama de Bode de um sistema $G(j\omega)$ determinado experimentalmente é indicado na Fig. 9.100. Determinar a função de transferência $G(s)$.

10

Técnicas de Projeto e Compensação

10.1 INTRODUÇÃO

O objetivo principal deste capítulo é apresentar procedimentos para o projeto e compensação de sistemas de controle invariantes no tempo, lineares, com entrada-simples-saída-simples. A compensação é o ajuste de um sistema a fim de satisfazer as especificações exigidas. As abordagens para projeto e compensação de sistemas de controle utilizadas neste capítulo são a abordagem do lugar das raízes e a abordagem de resposta em freqüência. (Será apresentada no Cap. 16 a teoria de controle moderna aplicada a projetos de sistemas.)

Especificações de desempenho. Os sistemas de controle são projetados para desempenhar tarefas específicas. Os requisitos impostos aos sistemas de controle normalmente são designados como especificações de desempenho. Geralmente são relativos à precisão, estabilidade relativa e velocidade de resposta.

Para problemas de projeto de rotina, as especificações de desempenho devem ser fornecidas em termos de valores numéricos precisos. Em outros casos, podem ser dadas parcialmente em termos de valores numéricos precisos e, parcialmente, em termos de hipóteses qualitativas. Neste último caso, as especificações podem ser modificadas durante o desenvolvimento do projeto, já que as especificações solicitadas podem não ser possíveis de satisfazer (devido a requisitos conflitantes) ou resultar em um sistema muito caro.

De um modo geral, as especificações de desempenho não devem ser mais restritivas do que o necessário para desempenhar uma dada tarefa. Se a precisão em operação em regime estacionário for a principal exigência de um dado sistema de controle, então não deverão ser exigidas especificações de desempenho desnecessariamente rígidas em relação à resposta transitória, desde que estas especificações exigirão componentes caros. Lembre-se que a parte mais importante do projeto de um sistema de controle é estabelecer precisamente as especificações de desempenho de modo que resultem em um sistema de controle ótimo para a finalidade requerida.

Abordagem de tentativa-e-erro para projeto de sistemas. Na maioria dos casos práticos, o método de projeto a ser utilizado pode ser determinado pelas especificações de desempenho aplicáveis ao caso particular. No projeto de sistemas de controle, se as especificações de desempenho forem dadas em termos de medidas de desempenho no domínio do tempo, tais como tempo de subida, sobrelevação máxima, ou tempo de estabelecimento, ou medidas de desempenho no domínio da freqüência, tais como margem de fase, margem de ganho, valor do pico de ressonância ou largura de faixa, então não temos escolha senão utilizar uma abordagem de tentativa-e-erro baseada no método do lugar das raízes e/ou nos métodos de resposta em freqüência.

Os sistemas que devem ser projetados através de uma abordagem de tentativa-e-erro normalmente são limitados a sistemas invariantes no tempo, lineares, com entrada-simples-saída-simples. O projetista deve tentar satisfazer todas as especificações de desempenho por meio de uma repetição educada de tentativa-e-erro. Após ser projetado o sistema, o projetista verifica se o sistema projetado satisfaz todas as especificações de desempenho. Se não for o caso, então repete o processo de projeto por ajuste de parâmetros ajustáveis, ou modificando a configuração do sistema, até atingir as especificações dadas. Embora o projeto seja baseado em um procedimento de tentativa-e-erro, a vivência e o conhecimento do projetista desempenham um papel importante em um projeto bem-sucedido. Um projetista experiente deve estar apto a projetar um sistema aceitável sem usar muitas tentativas.

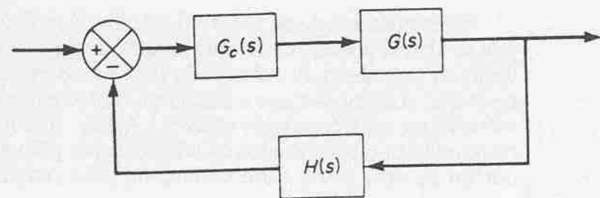
Modificação na dinâmica da planta (processo). Na construção de um sistema de controle, sabemos que modificações apropriadas na dinâmica do processo podem constituir um modo simples de atingir as especificações de desempenho. Isto, entretanto, pode não ser possível em muitas situações práticas porque o processo pode ser fixo ou não tolerar modificações. Então devemos ajustar outros parâmetros que não aqueles do processo fixado. Neste capítulo, suporemos que o processo é dado e inalterável.

Compensação de sistemas. O ajuste do ganho é o primeiro passo no ajuste de um sistema a fim de satisfazer um dado desempenho. Em muitos casos práticos, entretanto, o ajuste apenas do ganho pode não resultar em uma alteração suficiente no comportamento do sistema de modo a atingir as especificações dadas. Freqüentemente, aumentando o valor do ganho melhora-se o comportamento em regime estacionário, porém resulta uma estabilidade pobre ou mesmo uma instabilidade. É então necessário reprojetar o sistema (por modificação na estrutura ou incorporando dispositivos ou componentes adicionais) de modo a alterar o comportamento global para que o sistema se comporte como desejado.

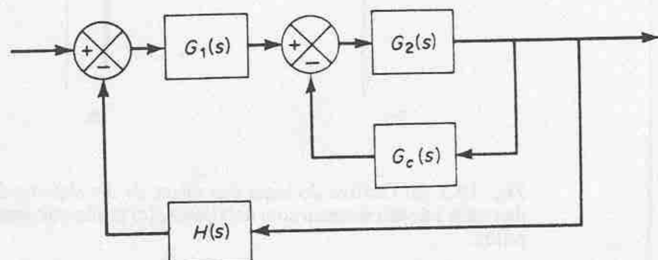
Um dispositivo adicional inserido no sistema para este propósito é denominado um *compensador*. Este dispositivo compensa o desempenho deficiente do sistema original.

Compensação série e compensação com realimentação (ou paralela). Se o compensador $G_c(s)$ é colocado em série com a função de transferência inalterável $G(s)$ como indicado na Fig. 10.1(a), então a compensação é denominada compensação série.

Uma alternativa para a compensação série é realimentar o(s) sinal(is) de algum(ns) elemento(s) e colocar um compensador no caminho de realimentação interna resultante, conforme indicado na Fig. 10.1(b). Esta compensação é denominada compensação por realimentação ou compensação paralela.



(a)



(b)

Fig. 10.1 (a) Compensação série; (b) compensação por realimentação ou paralela.

Em sistemas de controle compensados, verificamos que o problema normalmente se resume a um projeto conveniente de um compensador série ou por realimentação. A escolha entre a compensação série ou a compensação por realimentação depende da natureza dos sinais no sistema, dos níveis de potência em vários pontos, de componentes disponíveis, da experiência do projetista, de considerações econômicas etc.

Em geral, a compensação série pode ser mais simples do que a compensação por realimentação; entretanto, a compensação série normalmente exige amplificadores adicionais para aumentar o ganho e/ou propiciar isolação. (Para evitar dissipação de potência, o compensador série é inserido no ponto de menor energia do ramo direto.) Note que, em geral, o número de componentes necessários na compensação por realimentação é menor do que o número de componentes para a compensação série, desde que seja disponível um sinal conveniente, porque a transferência de energia é de um nível de maior potência para um outro de menor potência. (Isto significa que eventualmente não são necessários amplificadores adicionais.)

Compensadores. Se for necessário um compensador para atingir as especificações de desempenho, o projetista deve implementar um dispositivo físico que possua a função de transferência prescrita do compensador.

Vários dispositivos físicos são utilizados para estes fins. De fato, muitas idéias excelentes e úteis para implementações físicas de compensadores podem ser encontradas na literatura.

Dentre os muitos compensadores, os compensadores-série mais empregados são os chamados compensadores em avanço, compensadores em atraso e compensadores avanço-atraso. Neste capítulo, limitaremos nossa discussão exclusivamente a estes três tipos. Normalmente são elétricos, mecânicos, pneumáticos, hidráulicos ou combinações entre estes e consistem em redes *RC* (elétrica, mecânica, pneumática ou hidráulica) e amplificadores. (Em muitos casos, uma rede *RC* compensadora constitui parte integral do amplificador.)

Em projeto real de um sistema de controle, a escolha do uso de um compensador elétrico, mecânico, pneumático ou hidráulico é algo que deve ser decidido parcialmente em função da natureza da planta controlada. Por exemplo, se a planta controlada envolve um fluido inflamável, então temos de escolher um compensador e atuador pneumático a fim de evitar a possibilidade de faíscas. Se, entretanto, não existe o perigo de incêndio, então são mais comumente usados os compensadores eletrônicos. (De fato, normalmente os sinais não elétricos são transformados em sinais elétricos devido à simplicidade de transmissão, aumento de precisão, aumento de confiabilidade, facilidade de compensação etc.)

Procedimentos de projeto. Na abordagem de tentativa-e-erro para projeto de um sistema, definimos um modelo matemático do sistema de controle e ajustamos os parâmetros de um compensador. A parte que mais despende tempo neste trabalho é a verificação das especificações de desempenho pela análise com cada ajuste dos parâmetros, e o projetista pode necessitar fazer uso de um computador analógico ou digital para evitar o trabalho numérico maçante para esta verificação.

Uma vez obtido um modelo matemático satisfatório, o projetista deve construir um protótipo e testar o sistema de malha-aberta. Se estivêr assegurada a estabilidade absoluta, o projetista fecha a malha e verifica o desempenho do sistema na malha-fechada resultante. Devido aos efeitos de carga entre os componentes que foram desprezados, não linearidade, parâmetros distribuídos etc., que também não foram considerados no trabalho do projeto original, o desempenho real do sistema protótipo provavelmente difere das previsões teóricas. Portanto, o primeiro projeto pode não satisfazer todas as exigências de desempenho, e por tantativa-e-erro, o projetista deve efetuar modificações no protótipo até que o sistema atinja as especificações. Para fazer isto ele deve analisar cada tentativa, e os resultados da análise devem ser incorporados na próxima tentativa. O projetista deve verificar que o sistema final satisfaz as especificações de desempenho e, ao mesmo tempo, é confiável e econômico.

É importante notar que em um projeto realizado através da abordagem de tentativa-e-erro, ou em um projeto através de análise das especificações dadas, não resultará um sistema único. De fato, muitos (e possivelmente um número infinito) sistemas podem satisfazer as especificações fornecidas. Uma escolha ótima entre as muitas possibilidades pode ser feita a partir de considerações como desempenho global projetado, custo, dimensões e peso.

Projeto de sistemas complexos. As abordagens do lugar das raízes e de resposta em freqüência para projetos que consistem essencialmente em ajuste do ganho e do projeto de compensadores são muito úteis porém limitadas a sistemas de controle ideais e relativamente simples, tais como sistemas invariantes no tempo, lineares, com entrada-simples-saída-simples. Estas abordagens de projeto possuem severas limitações e dificuldades quando aplicadas a projetos de sistemas variáveis no tempo e com entradas-múltiplas-saídas-múltiplas.

Enquanto o projeto de sistemas de controle através das abordagens do lugar das raízes e de resposta em freqüência constitui um esforço de engenharia, o projeto de sistemas no contexto da teoria de controle moderna (a ser apresentada no Cap. 16) emprega formulações matemáticas do problema e aplica teoria matemática para projetar problemas nos quais os sistemas podem possuir múltiplas entradas e saídas e ser variáveis no tempo. Aplicando a teoria de controle moderna, o projetista está apto a iniciar de um índice de desempenho, juntamente com os vínculos impostos no sistema, e proceder a um projeto de um sistema estável através de um procedimento completamente analítico. A vantagem do projeto baseado nesta teoria de controle é que ela possibilita ao projetista produzir um sistema de controle que é ótimo em relação ao índice de desempenho considerado.

É importante ressaltar, entretanto, que esta técnica de projeto não pode ser aplicada se as especificações de desempenho forem dadas em termos de dados no domínio do tempo ou no domínio de freqüência, quando então as técnicas do lugar das raízes e de resposta em freqüência resultam mais úteis.

10.2 CONSIDERAÇÕES PRELIMINARES DE PROJETO

Os problemas de projeto considerados neste capítulo são aqueles que correspondem a uma melhora no desempenho de um sistema pela inserção de um compensador. A compensação de um sistema de controle é restrita ao projeto de um filtro cujas características tendem a compensar as características indesejáveis e inalteráveis do processo.

Nas Seções 10.3 até 10.5, consideraremos especificamente o projeto de compensadores em avanço, compensadores em atraso e compensadores avanço-atraso. Nestes problemas de projeto, colocamos um compensador em série com uma função de transferência inalterável $G(s)$ de modo a obter o comportamento desejável. O principal problema constitui então na escolha criteriosa do(s) pólo(s) e zero(s) do compensador $G_c(s)$ de modo a alterar o lugar das raízes ou a resposta em freqüência a fim de que sejam satisfeitas as especificações de desempenho.

Abordagem do lugar das raízes para projetos de sistemas de controle. O método do lugar das raízes é um método gráfico para a determinação das localizações de todos os pólos de malha-fechada a partir do conhecimento das localizações dos pólos e zeros de malha-aberta conforme algum parâmetro (normalmente o ganho) é variado desde zero até infinito. O método fornece uma clara indicação dos efeitos do ajuste do parâmetro. Uma vantagem do método do lugar das raízes é que verificamos ser possível obter informação tanto da resposta transitória como da resposta em freqüência a partir da configuração dos pólos e zeros do sistema no plano s.

Na prática, o gráfico do lugar das raízes de um sistema pode indicar que o desempenho desejado não pode ser conseguido somente pelo ajuste do ganho. De fato, em alguns casos, o sistema pode não ser estável para todos os valores de ganho. É necessário então modificar os lugares das raízes a fim de satisfazer as especificações de desempenho.

No projeto de um sistema de controle, se for exigido um outro ajuste além do ganho, devemos modificar os lugares das raízes originais através da inserção de um compensador conveniente. Uma vez entendidos completamente os efeitos sobre o lugar das raízes da adição de pólos e/ou zeros, podemos realmente determinar as localizações do(s) pólo(s) e do(s) zero(s) do compensador que modifique o lugar das raízes conforme desejado.

Efeitos da adição de pólos. A adição de um pólo na função de transferência de malha-aberta possui o efeito de puxar o lugar das raízes para a direita, tendendo a diminuir a estabilidade relativa do sistema e diminuindo o tempo de acomodação da resposta. (Lembre-se que a adição do controle integral adiciona um pólo na origem, tornando o sistema menos estável.) A Fig. 10.2 mostra exemplos dos lugares das raízes, ilustrando os efeitos da adição de um pólo a um sistema de pólos simples e a adição de dois pólos a um sistema de pólo simples.

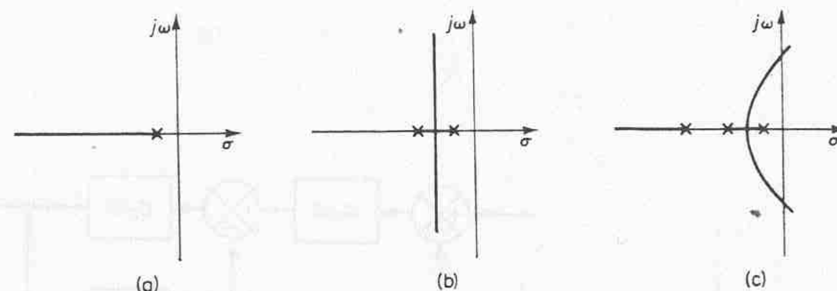


Fig. 10.2 (a) Gráfico do lugar das raízes de um sistema de pólo simples; (b) gráfico do lugar das raízes de um sistema com dois pólos; (c) gráfico do lugar das raízes de um sistema com três pólos.

Efeitos da adição de zeros. A adição de um zero na função de transferência de malha-aberta tem o efeito de puxar o lugar das raízes para a esquerda, tendendo a fazer o sistema mais estável e diminuindo o tempo de acomodação da resposta (fisicamente, a adição de um zero na função de transferência direta significa a adição de um controle derivativo ao sistema. O efeito deste controle é introduzir um grau de antecipação no sistema e aumentar a velocidade da resposta transitória). A Fig. 10.3(a) mostra os lugares das raízes para um sistema que é estável para pequenos ganhos mas instável para grandes ganhos. As Figs. 10.3(b), (c) e (d) mostram os gráficos dos lugares das raízes para o sistema quando um zero é adicionado à função de transferência de malha-aberta. Note que quando um zero é adicionado ao sistema da Fig. 10.3(a), este torna-se estável para todos os valores de ganho.

Abordagem de resposta em freqüência para projeto de sistemas de controle. Considerando o problema de sistemas de controle compensados através das técnicas no domínio de freqüência, asseguramos o controle do comportamento da resposta transitória em termos das especificações no domínio de freqüência, tais como margem de fase, margem de ganho, valor de pico de ressonância e largura de faixa (banda). O projeto no domínio de freqüência é indireto porque o sistema é projetado para satisfazer estas especificações no domínio de freqüência ao invés das especificações no domínio do tempo. Após haver projetado a malha-aberta pelo método da resposta em freqüência, podem ser obtidos os pólos e zeros de malha-fechada. Devem-se então verificar se as características da resposta transitória satisfazem ou não as especificações exigidas no domínio do tempo. Em caso negativo, o compensador deve ser modificado e a análise repetida até ser obtido um resultado satisfatório.

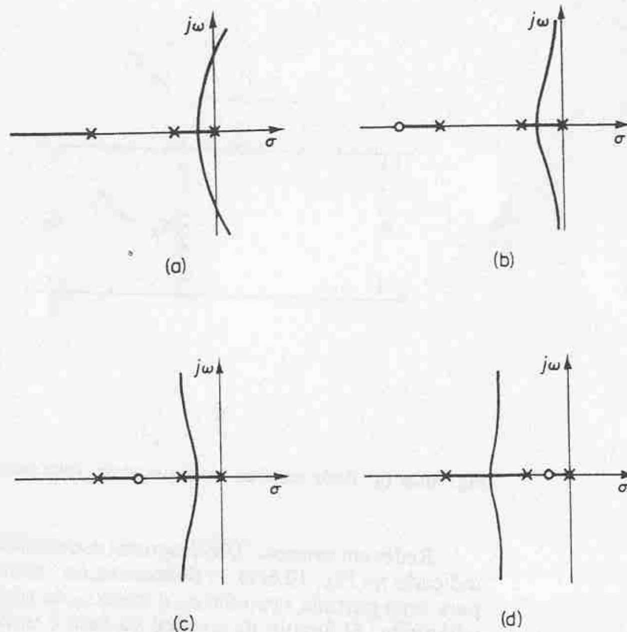


Fig. 10.3 (a) Gráfico do lugar das raízes de um sistema com três pólos; (b), (c) e (d) gráficos do lugar das raízes mostrando os efeitos da adição de um zero ao sistema com três pólos.

O projeto no domínio de freqüências é simples e direto. O gráfico da resposta em freqüência indica claramente a maneira pela qual o sistema deve ser modificado, embora não possa ser feita a previsão exata quantitativa das características da resposta transitória. A abordagem da resposta em freqüência pode ser aplicada a sistemas ou componentes cujas características dinâmicas são fornecidas na forma de dados de resposta em freqüência. Note que devido à dificuldade na obtenção das equações que governam alguns componentes, tais como componentes pneumáticos e hidráulicos, as características dinâmicas de tais componentes normalmente são determinadas experimentalmente através de testes de resposta em freqüência. Os gráficos de resposta em freqüência obtidos experimentalmente podem ser facilmente combinados com outros gráficos. Note também que considerando ruídos de alta freqüência, verificamos que a abordagem de resposta em freqüência é mais conveniente do que outras abordagens.

No projeto de sistemas de controle no domínio de freqüência, se desejarmos uma certa margem de fase ou margem de ganho, notamos que os diagramas de Bode são mais convenientes do que os diagramas polares. (No uso de diagramas de Bode, exceto nos pontos próximos à freqüência do cruzamento do ganho onde as curvas exatas diferem consideravelmente das assíntotas linearizadas, podemos empregar os gráficos com assíntotas para fins de projeto.) Por outro lado, se desejarmos um certo valor de M_r, os gráficos polares ou os gráficos log-módulo *versus* fase apresentam uso mais conveniente do que os diagramas de Bode.

Informação obtenível da resposta em freqüência de malha-aberta. A região de baixa freqüência (a região abaixo da freqüência de cruzamento do ganho —* eventualmente, freqüência de corte) indica o comportamento em regime estacionário do sistema em malha-fechada. A região de média freqüência (a região próxima ao ponto $-1 + j0$) do lugar geométrico, indica a estabilidade relativa. A região de alta freqüência (a região acima da freqüência de cruzamento do ganho) indica a complexidade do sistema.

Requisitos da resposta em freqüência em malha-aberta. Podemos dizer que, em muitos casos práticos, a compensação é essencialmente um compromisso entre o erro em regime estacionário e a estabilidade relativa.

A fim de possuir um alto valor do coeficiente de erro de velocidade e ainda uma estabilidade relativa satisfatória, verificamos ser necessário modificar a curva de resposta em freqüência de malha-aberta.

O ganho na região de baixa freqüência deve ser suficientemente grande, e, também, próxima à freqüência de cruzamento do ganho, a inclinação da curva do módulo no diagrama de Bode deve ser -20 db/década. Esta inclinação deve estender-se por uma faixa de freqüência suficientemente larga para assegurar uma

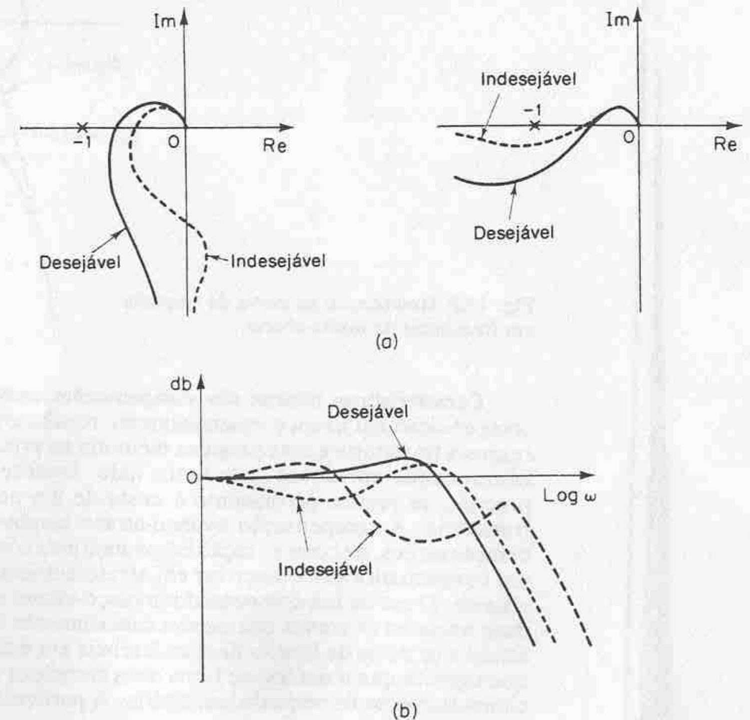


Fig. 10.4 (a) Exemplos de curvas de resposta em freqüência de malha-aberta desejáveis e indesejáveis; (b) exemplos de curvas de resposta em freqüência de malha-fechada desejáveis e indesejáveis.

*N. do T.

margem de fase apropriada. Para a região de alta freqüência, o ganho deve ser atenuado tão rapidamente quanto possível para minimizar os efeitos do ruído.

Exemplos de curvas de resposta em freqüência de malha-aberta e de malha-fechada desejáveis e indesejáveis são fornecidos na Fig. 10.4.

Em relação à Fig. 10.5 vemos que a modificação da curva de resposta em freqüência de malha-aberta pode ser feita se a parte de alta freqüência do lugar geométrico segue o lugar geométrico de $G_1(j\omega)$, enquanto a parte de baixa freqüência do lugar geométrico segue o lugar geométrico de $G_2(j\omega)$. O lugar geométrico modificado de $G_c(j\omega)G(j\omega)$ deve apresentar margens de fase e de ganho razoáveis ou ser tangente a um círculo M apropriado, conforme indicado.

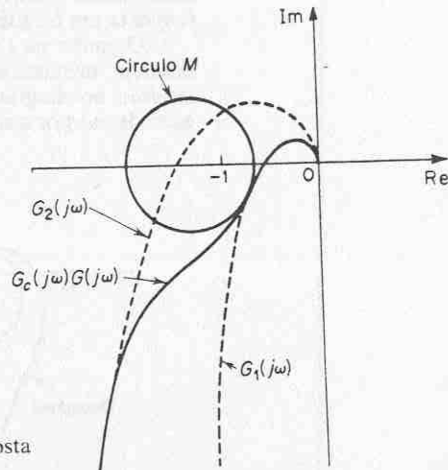


Fig. 10.5 Modificação na curva de resposta em freqüência de malha-aberta.

Características básicas das compensações avanço, atraso, avanço-atraso. A compensação em avanço essencialmente resulta em uma melhoria apreciável na resposta transitória e uma pequena melhoria na precisão de regime permanente. A compensação em atraso, por outro lado, fornece uma apreciável melhoria na precisão de regime permanente à custa de um aumento no tempo de resposta transitória. A compensação avanço-atraso combina as características das duas compensações, da compensação em avanço e da compensação em atraso. O uso de um compensador em avanço ou em atraso aumenta de uma unidade a ordem do sistema. O uso de um compensador avanço-atraso aumenta a ordem do sistema de duas unidades (a menos que ocorra cancelamento entre os zeros da rede avanço-atraso e os pólos da função de transferência em malha-aberta não compensada), o que significa que o sistema se torna mais complexo e mais difícil de controlar o seu comportamento na resposta transitória. A particular situação determina o tipo da compensação a ser utilizada.

10.3 COMPENSAÇÃO EM AVANÇO

Nesta seção, deduziremos inicialmente as funções de transferência de uma rede em avanço elétrica e de uma rede mecânica em avanço. Apresentaremos, então, procedimentos para projetar compensadores em avanço baseados nas abordagens do lugar das raízes e da resposta em freqüência.

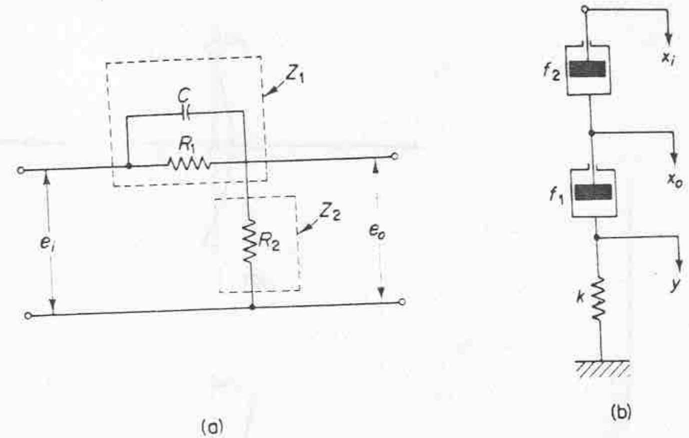


Fig. 10.6 (a) Rede elétrica em avanço; (b) rede mecânica em avanço.

Redes em avanço. Um diagrama esquemático de uma rede elétrica em avanço é indicado na Fig. 10.6(a). A denominação "rede em avanço" resulta do fato de que para uma entrada senoidal e_i, a saída e_0 da rede também é senoidal com uma fase adiantada. O ângulo de avanço da fase é uma função da freqüência de entrada. Vamos deduzir a função de transferência para esta rede. Conforme é usual na dedução da função de transferência de qualquer rede de quatro terminais, suporemos que a impedância da fonte vista pela rede é nula e que a impedância de carga da saída é infinita.

Usando os símbolos definidos na Fig. 10.6(a), verificamos que as impedâncias complexas Z_1 e Z_2 são

$$Z_1 = \frac{R_1}{R_1Cs + 1}, \qquad Z_2 = R_2$$

A função de transferência entre a saída $E_0(s)$ e a entrada $E_i(s)$ é

$$\frac{E_0(s)}{E_i(s)} = \frac{Z_2}{Z_1 + Z_2} = \frac{R_2}{R_1 + R_2}\,\frac{R_1Cs + 1}{\dfrac{R_1R_2}{R_1 + R_2}Cs + 1}$$

Definindo,

$$R_1C = T, \qquad \frac{R_2}{R_1 + R_2} = \alpha < 1$$

Então a função de transferência resulta em

$$\frac{E_0(s)}{E_i(s)} = \alpha\,\frac{Ts + 1}{\alpha Ts + 1} = \frac{s + \dfrac{1}{T}}{s + \dfrac{1}{\alpha T}}$$

A Fig. 10.6(b) mostra um diagrama esquemático de uma rede mecânica de avanço. Do diagrama, obtemos as seguintes equações:

$$f_2(\dot{x}_i - \dot{x}_0) = f_1(\dot{x}_0 - \dot{y})$$

$$f_1(\dot{x}_0 - \dot{y}) = ky$$

Calculando as transformadas de Laplace destas duas equações, supondo condições iniciais nulas, e então eliminando $Y(s)$, obtemos

$$\frac{X_0(s)}{X_i(s)} = \frac{f_2}{f_1 + f_2}\,\frac{\frac{f_1}{k}s + 1}{\frac{f_2}{f_1 + f_2}\frac{f_1}{k}s + 1}$$

Esta é a função de transferência entre $X_0(s)$ e $X_i(s)$. Por definição,

$$\frac{f_1}{k} = T, \qquad \frac{f_2}{f_1 + f_2} = \alpha < 1$$

obtemos

$$\frac{X_0(s)}{X_i(s)} = \alpha\frac{Ts + 1}{\alpha Ts + 1} = \frac{s + \frac{1}{T}}{s + \frac{1}{\alpha T}}$$

Características de redes em avanço. Uma rede em avanço possui a seguinte função de transferência:

$$\alpha\frac{Ts + 1}{\alpha Ts + 1} = \frac{s + \frac{1}{T}}{s + \frac{1}{\alpha T}} \qquad (\alpha < 1)$$

A função de transferência possui um zero em $s = -1/T$ e um pólo em $s = -1/(\alpha T)$. Desde que $\alpha < 1$, verificamos que o zero está sempre localizado à direita do pólo no plano complexo. Note que para um valor pequeno de α, o pólo está localizado à esquerda e muito distante. O valor mínimo de α é limitado pela implementação física da rede em avanço. O valor mínimo de α normalmente é considerado aproximadamente 0,07. Se o valor de α é pequeno, é necessário colocar em cascata um amplificador a fim de compensar a atenuação da rede em avanço.

A Fig. 10.7 indica o gráfico polar de

$$\alpha\frac{j\omega T + 1}{j\omega\alpha T + 1} \qquad (0 < \alpha < 1)$$

Para um dado valor de α, o ângulo entre o eixo real positivo e a reta tangente desenhada a partir da origem ao semicírculo fornece o ângulo de avanço de fase

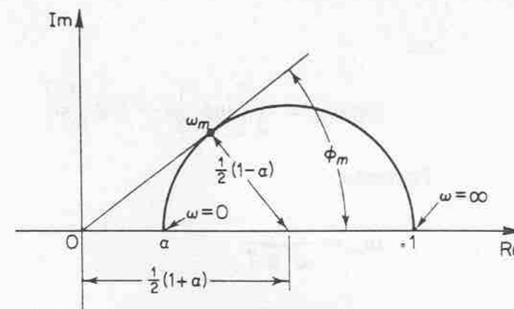


Fig. 10.7 Diagrama polar de uma rede em avanço $\alpha(j\omega T + 1)/(j\omega\alpha T + 1)$, onde $0 < \alpha < 1$.

máximo, ϕ_m. Denominaremos a freqüência ω_m no ponto tangente. Da Fig. 10.7, o ângulo de fase em $\omega = \omega_m$ é

$$\operatorname{sen}\phi_m = \frac{\frac{1-\alpha}{2}}{\frac{1+\alpha}{2}} = \frac{1-\alpha}{1+\alpha} \tag{10.1}$$

A Eq. (10.1) relaciona o ângulo de avanço de fase máximo com o valor de α.

A Fig. 10.8 mostra o diagrama de Bode de uma rede em avanço quando $\alpha = 0.1$. As freqüências de canto para a rede em avanço são $\omega = 1/T$ e $\omega = 1/(\alpha T)$. Examinando a Fig. 10.8, verificamos que ω_m é a média geométrica das duas freqüências de canto.

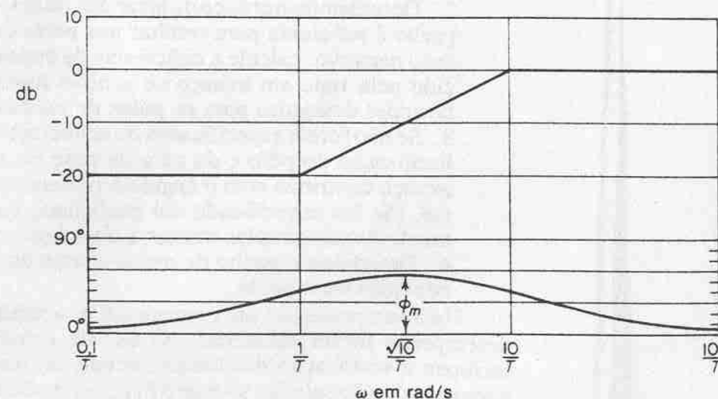


Fig. 10.8 Diagrama de Bode de uma rede em avanço.

ou

$$\log \omega_m = \frac{1}{2}\left[\log \frac{1}{T} + \log \frac{1}{\alpha T}\right]$$

Portanto,

$$\omega_m = \frac{1}{\sqrt{\alpha}\, T} \tag{10.2}$$

Conforme pode ser visto na Fig. 10.8, a rede em avanço é basicamente um filtro passa-altas. (Passam as altas freqüências porém são atenuadas as baixas freqüências.) Portanto, é necessário um ganho adicional, eventualmente, para aumentar o ganho em baixas freqüências.

Técnicas de compensação em avanço baseadas na abordagem do lugar das raízes. A abordagem do lugar das raízes para projetos é muito eficiente quando as especificações são fornecidas em termos de grandezas no domínio do tempo, tais como sobrelevação máxima, tempo de subida, tempo de acomodação, relação de amortecimento, e freqüência natural não amortecida dos pólos de malha-fechada dominantes desejados.

Considere um problema de projeto no qual o sistema original ou é instável para todos os valores de ganho ou é estável porém possui características indesejáveis de resposta transitória. Neste caso, é necessária uma modificação no lugar das raízes na vizinhança do eixo $j\omega$ e da origem, de modo que os pólos de malha-fechada dominantes estejam nas posições desejadas do plano complexo. Este problema pode ser resolvido pela inserção de um compensador em avanço apropriado em cascata com a função de transferência do ramo direto.

Os procedimentos para projeto de um compensador em avanço pelo método do lugar das raízes podem ser estabelecidos como segue:

1. A partir das especificações de desempenho, determine as localizações desejadas para os pólos de malha-fechada dominantes.
2. Desenhando o gráfico do lugar das raízes, verifique se apenas um ajuste do ganho é suficiente para resultar nos pólos de malha-fechada desejados. Em caso negativo, calcule a deficiência de ângulo ϕ. Este ângulo deve ser fornecido pela rede em avanço se o novo lugar das raízes passar através das posições desejadas para os pólos de malha-fechada dominantes.
3. Se não forem especificados os coeficientes de erros estáticos, determine a localização do pólo e do zero da rede em avanço de modo que a rede em avanço contribua com o ângulo ϕ necessário e exija o mínimo ganho adicional. (Se for especificado um coeficiente particular de erro estático, geralmente é mais simples utilizar a abordagem da resposta em freqüência.)
4. Determine o ganho de malha-aberta do sistema compensado através da condição do módulo.

Uma vez projetado um compensador, verifique se todas as especificações de desempenho foram satisfeitas. (O uso de computadores analógicos ou digitais facilitam a verificação das características da resposta transitória.) Se o sistema compensado não satisfaz as especificações de desempenho, repita o procedimento de projeto ajustando o pólo e o zero do compensador até todas as especificações serem satisfeitas. Se for exigido um coeficiente de erro estático grande, coloque em cascata uma rede em atraso ou altere o compensador em avanço para um compensador avanço-atraso.

Note que se os pólos de malha-fechada dominantes selecionados não forem realmente dominantes, será necessário modificar a localização do par de pólos de malha-fechada através de uma abordagem de tentativa-e-erro. (Os pólos de malha-fechada que não são dominantes modificam apenas a resposta obtida dos pólos de malha-fechada. A quantidade de modificação depende da posição destes pólos de malha-fechada restantes.)

Exemplo 10.1 Considere o sistema indicado na Fig. 10.9(a). A função de transferência do ramo direto é

$$G(s) = \frac{4}{s(s+2)}$$

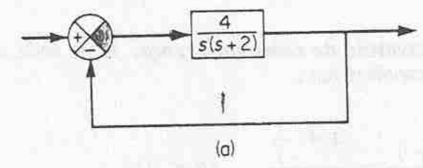
O gráfico do lugar das raízes para este sistema é indicado na Fig. 10.9(b). A função de transferência de malha-fechada resulta em

$$\frac{C(s)}{R(s)} = \frac{4}{s^2 + 2s + 4}$$

$$= \frac{4}{(s + 1 + j\sqrt{3})(s + 1 - j\sqrt{3})}$$

Os pólos de malha-fechada estão localizados em

$$s = -1 \pm j\sqrt{3}$$



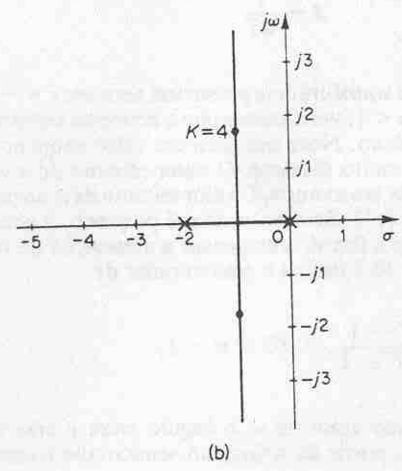


Fig. 10.9 (a) Sistema de controle: (b) gráfico do lugar das raízes.

A relação de amortecimento dos pólos de malha-fechada é 0,5. A freqüência natural não amortecida dos pólos de malha-fechada é 2 rad/s. O coeficiente de erro de velocidade estático é 2 s^{-1}.

Deseja-se modificar os pólos de malha-fechada de modo que uma freqüência natural não amortecida $\omega_n = 4$ rad/s seja obtida sem modificar o valor da relação de amortecimento, $\zeta = 0,5$.

Lembre-se que a relação de amortecimento ζ de um par de pólos complexos no plano complexo pode ser expressa em termos do ângulo θ que é medido a partir do eixo real negativo, conforme indicado na Fig. 10.10(a), com

$$\zeta = \cos\theta$$

Em outras palavras, retas com relação de amortecimento ζ constantes são retas radiais passando pela origem conforme a Fig. 10.10(b). Por exemplo, uma relação de amortecimento de 0,5 exige que os pólos complexos se situem sobre as retas que passam pela origem e fazem ângulos de $\pm 60°$ com o eixo real negativo. (Se a parte real de um par de pólos complexos é positiva, significando que o sistema é instável, o ζ correspondente é negativo.) A relação de amortecimento determina a posição angular do pólo, enquanto a distância do pólo em relação à origem é determinada pela freqüência natural não amortecida ω_n.

No presente exemplo, as posições desejadas dos pólos de malha-fechada são

$$s = -2 \pm j2\sqrt{3}$$

Em alguns casos, após haverem sido obtidos os lugares das raízes do sistema original, os pólos de malha-fechada dominantes podem ser movidos para as posições desejadas através de um simples ajuste do ganho. Este não é o caso, entretanto, para o sistema apresentado. Portanto, inserimos um compensador em avanço no ramo direto.

Um procedimento geral para determinação do compensador em avanço é o seguinte: Inicialmente, determine a soma dos ângulos na posição desejada de um dos pólos de malha-fechada dominante com os pólos e zeros de malha-aberta do sistema original e o ângulo ϕ

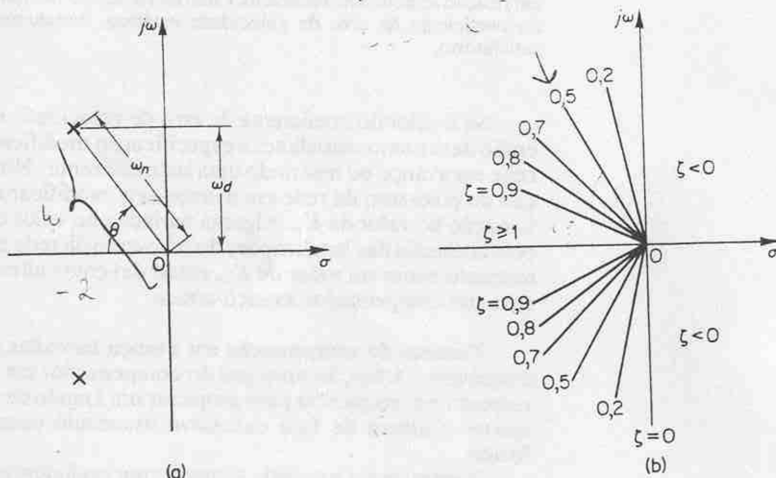


Fig. 10.10 (a) Pólos complexos; (b) retas de relação de amortecimento ζ constante.

necessário para ser adicionado de modo que a soma total dos ângulos seja igual a $\pm 180°(2k + 1)$. A rede em avanço deve contribuir com este ângulo. (Se o ângulo for muito grande, então duas ou mais redes em avanço podem ser necessárias em vez de apenas uma.)

Se o sistema original possui a função de transferência de malha-aberta $G(s)$, então o sistema compensado possuirá a função de transferência de malha-aberta

$$G_1(s) = \left(\alpha \frac{Ts+1}{\alpha Ts+1}\right) K_c G(s)$$

onde o primeiro termo do segundo membro da equação corresponde à rede em avanço, o segundo termo K_c é o ganho do amplificador, e o último termo $G(s)$ é a função de transferência de malha-aberta original. (Note que o amplificador possibilita o casamento de impedância desejado bem como o ganho K_c desejado.) Note que há muitos valores possíveis para T que resultarão na contribuição do ângulo necessário para os pólos de malha-fechada desejados.

O próximo passo é determinar as posições do pólo e zero da rede em avanço; em outras palavras, o valor de T. Na escolha do valor de T, apresentaremos um procedimento para obter o maior valor possível para α de modo que o ganho adicional exigido do amplificador seja o menor possível. Primeiro, desenhe uma reta horizontal passando pelo ponto P, a posição desejada de um dos pólos de malha-fechada dominantes. Isto é indicado pela reta PA na Fig. 10.11. Desenhe também uma reta ligando o ponto P à origem. Bisseccione o ângulo entre as retas PA e PO, conforme indicado na Fig. 10.11. Desenhe duas retas PC e PD que fazem ângulos de $\pm\phi/2$ com a bissetriz PB. As intersecções de PC e PD com o eixo real negativo fornecem as posições necessárias para o pólo e o zero da rede em avanço. O compensador assim projetado tornará o ponto P um ponto sobre o lugar das raízes do sistema compensado. O ganho de malha-aberta é determinado por meio da condição do módulo.

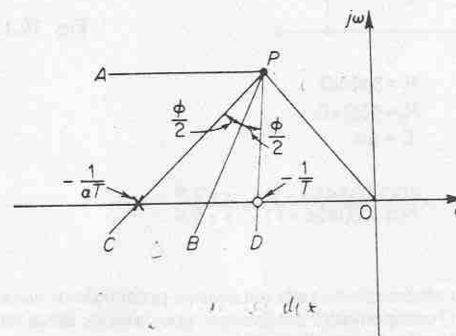


Fig. 10.11 Determinação do pólo e do zero de uma rede em avanço.

No presente sistema, o ângulo de $G(s)$ no pólo de malha-fechada desejado é

$$\left. \angle \frac{4}{s(s+2)} \right|_{s=-2+j2\sqrt{3}} = -210°$$

Portanto, para que o lugar das raízes passe através do pólo de malha-fechada desejado, a rede em avanço deve contribuir com $\phi = 30°$ neste ponto. Seguindo o procedimento de projeto anteriormente citado, determinamos o pólo e o zero da rede em avanço, conforme indicado na Fig. 10.12, obtendo

$$\text{Pólo em } s = -5,4, \qquad \text{Zero em } s = -2,9$$

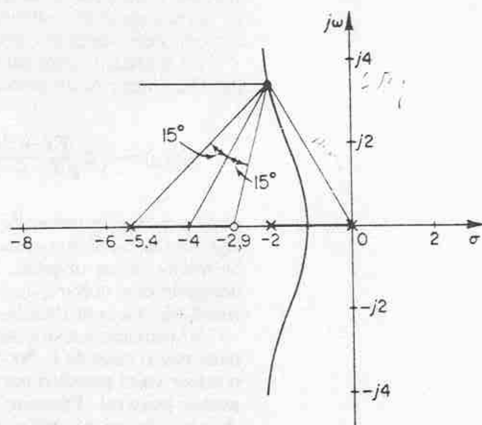


Fig. 10.12 Gráfico do lugar das raízes do sistema compensado.

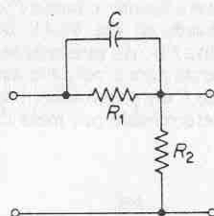


Fig. 10.13 Rede em avanço.

$R_1 = 345\ \text{k}\Omega$
$R_2 = 400\ \text{k}\Omega$
$C = 1\mu\text{f}$

Função de transferência $= \dfrac{400}{745}\left(\dfrac{0{,}345s+1}{0{,}185s+1}\right) = \dfrac{s+2{,}9}{s+5{,}4}$

Destes valores obtemos uma rede em avanço possuindo os valores dos parâmetros indicados na Fig. 10.13. O compensador em avanço, que consiste nesta rede de avanço e um amplificador, possui a função de transferência

$$G_c(s) = \frac{(s+2{,}9)}{(s+5{,}4)}K_c$$

Portanto, a função de transferência de malha-aberta do sistema compensado resulta em

$$G_c(s)G(s) = \frac{(s+2{,}9)}{(s+5{,}4)}K_c\frac{4}{s(s+2)} = \frac{K(s+2{,}9)}{s(s+2)(s+5{,}4)}$$

O gráfico do lugar das raízes para o sistema compensado é mostrado na Fig. 10.12. O ganho K é calculado a partir da condição do módulo como segue:

$$\left|\frac{K(s+2{,}9)}{s(s+2)(s+5{,}4)}\right|_{s=-2+j2\sqrt{3}} = 1$$

ou

$$K = 18{,}7$$

Segue que

$$G_c(s)G(s) = \frac{18{,}7(s+2{,}9)}{s(s+2)(s+5{,}4)}$$

A constante de ganho K_c do amplificador é

$$K_c = \frac{K}{4} = 4{,}68$$

O coeficiente de erro de velocidade estático K_v é obtido da expressão

$$\begin{aligned}K_v &= \lim_{s\to 0} sG_c(s)G(s)\\ &= \lim_{s\to 0}\frac{s18{,}7(s+2{,}9)}{s(s+2)(s+5{,}4)}\\ &= 5{,}02\ \text{s}^{-1}\end{aligned}$$

O terceiro pólo de malha-fechada é obtido dividindo-se a equação característica pelos fatores conhecidos como segue:

$$s(s+2)(s+5{,}4) + 18{,}7(s+2{,}9) = (s+2+j2\sqrt{3})(s+2-j2\sqrt{3})(s+3{,}4)$$

O método de compensação anteriormente apresentado possibilita colocar os pólos de malha-fechada dominantes nos pontos desejados do plano complexo. O terceiro pólo, em $s = -3{,}4$, está próximo do zero adicionado em $s = -2{,}9$. Portanto, o efeito deste pólo na resposta transitória é relativamente pequeno. Desde que não foi imposta nenhuma restrição em relação ao pólo não dominante e não foi fornecida nenhuma especificação relativa ao valor do coeficiente de erro de velocidade estático, concluímos que o projeto apresentado é satisfatório.

Se o valor do coeficiente de erro de velocidade estática K_v fosse especificado, então deveríamos satisfazer a especificação modificando a posição do pólo-zero da rede em avanço ou inserindo uma rede diferente. Note que modificando a localização do pólo-zero da rede em avanço sem modificar o ângulo ϕ, o resultado é uma variação no valor de K_v. Alguma variação no valor de K_v pode, portanto, ser feita pela alteração das localizações do pólo-zero da rede em avanço. Se for desejado um aumento maior no valor de K_v, então devemos alterar o compensador em avanço para um compensador avanço-atraso.

Técnicas de compensação em avanço baseadas na abordagem de resposta em freqüência. A função principal do compensador em avanço é modificar a curva de resposta em freqüência para propiciar um ângulo de avanço de fase suficiente para ajustar o atraso de fase excessivo associado com os componentes do sistema fixado.

Vamos supor um dado sistema com realimentação unitária. Desejamos satisfazer as especificações de desempenho, que são fornecidas em termos de margem de fase, margem de ganho, coeficientes de erro etc. Os procedimentos para projetar

um compensador em avanço, através da abordagem de resposta em freqüência, podem ser estabelecidos como segue:

1. Determine o ganho de malha-aberta K a fim de satisfazer as exigências dos coeficientes de erro.
2. Usando o ganho K anteriormente determinado, calcule a margem de fase do sistema não compensado.
3. Determine o ângulo de avanço de fase necessário ϕ a ser adicionado ao sistema.
4. Determine o fator de atenuação α pelo uso da Eq. (10.1). Determine a freqüência onde o módulo do sistema não compensado é igual a $-20 \log (1/\sqrt{\alpha})$. Selecione esta freqüência como a nova freqüência de cruzamento do ganho. Esta freqüência corresponde a ω_m e o deslocamento de fase máximo ϕ_m ocorre nesta freqüência.
5. Determine as freqüências de canto da rede em avanço a partir de

$$\omega = \frac{1}{T}, \qquad \omega = \frac{1}{\alpha T}$$

Finalmente, introduza um amplificador com ganho igual a $1/\alpha$, ou aumente o ganho do amplificador existente por um fator de $1/\alpha$.

Exemplo 10.2 Considere o sistema indicado na Fig. 10.14. A função de transferência de malha-aberta é

$$G(s) = \frac{4K}{s(s+2)}$$

Deseja-se determinar um compensador para o sistema de modo que o coeficiente de erro de velocidade estático K_v seja 20 s^{-1}, a margem de fase seja pelo menos 50°, e a margem de ganho pelo menos 10 db.

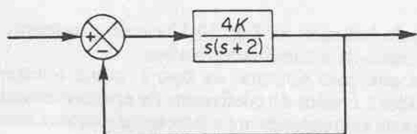


Fig. 10.14 Sistema de controle.

No presente exemplo, foram especificadas as margens de ganho e de fase; conseqüentemente, empregaremos os diagramas de Bode.

O primeiro passo no projeto é ajustar o ganho K para satisfazer a especificação de desempenho em regime estacionário, ou propiciar o coeficiente de erro de velocidade estático exigido. Desde que este coeficiente é dado como 20 s^{-1}, obtemos

$$K_v = \lim_{s \to 0} sG(s) = \lim_{s \to 0} \frac{s4K}{s(s+2)} = 2K = 20$$

ou

$$K = 10$$

Com $K = 10$, o sistema da Fig. 10.14 satisfaz a exigência de regime estacionário.

Desenharemos agora o gráfico do diagrama de Bode de

$$G(j\omega) = \frac{40}{j\omega(j\omega + 2)} = \frac{20}{j\omega(0{,}5\,j\omega + 1)}$$

A Fig. 10.15 mostra as curvas de módulo e ângulo de fase de $G(j\omega)$. Destes gráficos, as margens de fase e de ganho do sistema são obtidas como 17° e $+\infty$ db, respectivamente. (Uma margem de fase de 17° implica que o sistema é quase oscilatório. Portanto, satisfazendo apenas a especificação de regime estacionário, resulta em um desempenho pobre na resposta transitória.) A especificação exigida para a margem de fase é de pelo menos 50°. Portanto, determinaremos o ângulo de avanço de fase adicional necessário para satisfazer a exigência de estabilidade relativa, que é igual a 33°. Para conseguir uma margem de fase de 50° sem diminuir o valor de K, é necessário introduzir no sistema um conveniente compensador em avanço.

Observando que a adição de um compensador em avanço modifica a curva de módulo no diagrama de Bode, verificamos que a freqüência de cruzamento do ganho deverá ser deslocada para a direita. Devemos ajustar o ângulo de fase considerando um atraso de $G(j\omega)$ aumentado devido a este aumento na freqüência de cruzamento do ganho. Considerando o deslocamento da freqüência de cruzamento do ganho, devemos supor que ϕ_m, o ângulo de fase em avanço máximo necessário, é aproximadamente 38°. (Isto significa que 5° foram adicionados para compensar o deslocamento na freqüência de cruzamento do ganho.) Desde que

$$\operatorname{sen} \phi_m = \frac{1 - \alpha}{1 + \alpha}$$

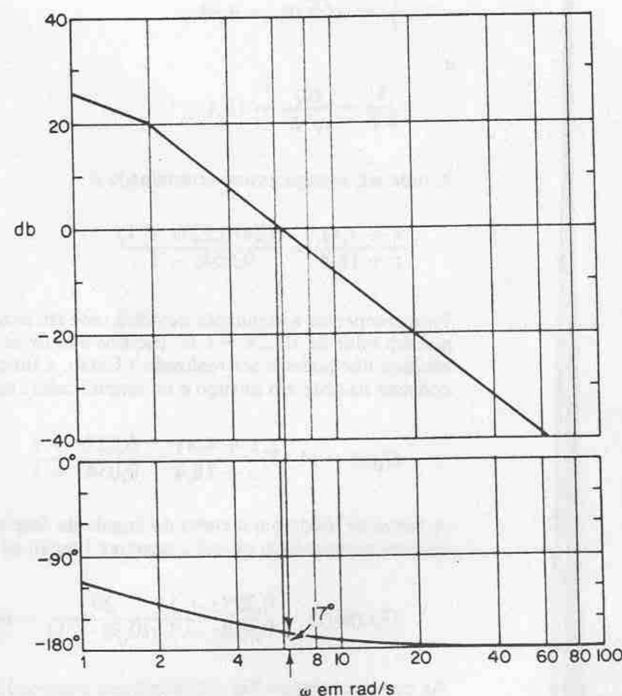


Fig. 10.15 Diagrama de Bode para $G(j\omega) = 40/[j\omega(j\omega + 2)]$.

$\phi_m = 38°$ corresponde a $\alpha = 0.24$. Uma vez que o fator de atenuação α foi determinado com base no ângulo de fase em avanço exigido, o próximo passo é determinar as freqüências de canto $\omega = 1/T$ e $\omega = 1/(\alpha T)$ da rede em avanço. Para assim proceder, note inicialmente que o ângulo de avanço de fase máximo ϕ_m ocorre na média geométrica das duas freqüências de canto, ou $\omega = 1/(\sqrt{\alpha}T)$. [Vide Eq. (10.2).] A quantidade de modificação na curva de módulo em $\omega = 1/(\sqrt{\alpha}T)$ é

$$\left|\frac{1 + j\omega T}{1 + j\omega\alpha T}\right|_{\omega = 1/(\sqrt{\alpha}T)} = \left|\frac{1 + j\dfrac{1}{\sqrt{\alpha}}}{1 + j\alpha\dfrac{1}{\sqrt{\alpha}}}\right| = \frac{1}{\sqrt{\alpha}}$$

Note que

$$\frac{1}{\sqrt{\alpha}} = \frac{1}{\sqrt{0{,}24}} = \frac{1}{0{,}49} = 6{,}2 \text{ db}$$

e $|G(j\omega)| = -6.2$ db corresponde a $\omega = 9$ rad/s. Selecionaremos esta freqüência como a nova freqüência de cruzamento do ganho ω_c. Observando que esta freqüência corresponde a $1/(\sqrt{\alpha}T)$, ou $\omega_c = 1/(\sqrt{\alpha}T)$, obtemos

$$\frac{1}{T} = \sqrt{\alpha}\,\omega_c = 4{,}41$$

e

$$\frac{1}{\alpha T} = \frac{\omega_c}{\sqrt{\alpha}} = 18{,}4$$

A rede em avanço assim determinada é

$$\frac{s + 4{,}41}{s + 18{,}4} = \frac{0{,}24(0{,}227s + 1)}{0{,}054s + 1}$$

Para compensar a atenuação devida à rede em avanço, aumentamos o ganho do amplificador por um fator de $1/0.24 = 4.17$. (Se isto não fosse feito, o coeficiente de erro de velocidade estático não poderia ser realizado.) Então, a função de transferência do compensador, que consiste na rede em avanço e no amplificador, resulta em

$$G_c(s) = (4{,}17)\frac{s + 4{,}41}{s + 18{,}4} = \frac{0{,}227s + 1}{0{,}054s + 1}$$

A curva de módulo e a curva do ângulo de fase para $G_c(j\omega)$ são indicadas na Fig. 10.16. O sistema compensado possui a seguinte função de transferência de malha-aberta:

$$G_c(s)G(s) = \frac{0{,}227s + 1}{0{,}054s + 1}\,\frac{20}{s(0{,}5s + 1)} = (4{,}17)\frac{s + 4{,}41}{s + 18{,}4}\,\frac{40}{s(s + 2)}$$

As curvas cheias na Fig. 10.16 indicam a curva do módulo e a curva do ângulo de fase para o sistema compensado. O compensador em avanço acarreta um aumento na freqüência de cruzamento do ganho de 6.3 para 9 rad/s. O aumento nesta freqüência significa um aumento na largura de faixa. Isto implica um aumento na velocidade de resposta. As margens de ganho e de fase são obtidas como aproximadamente 50° e $+\infty$ db, respectivamente. O sistema

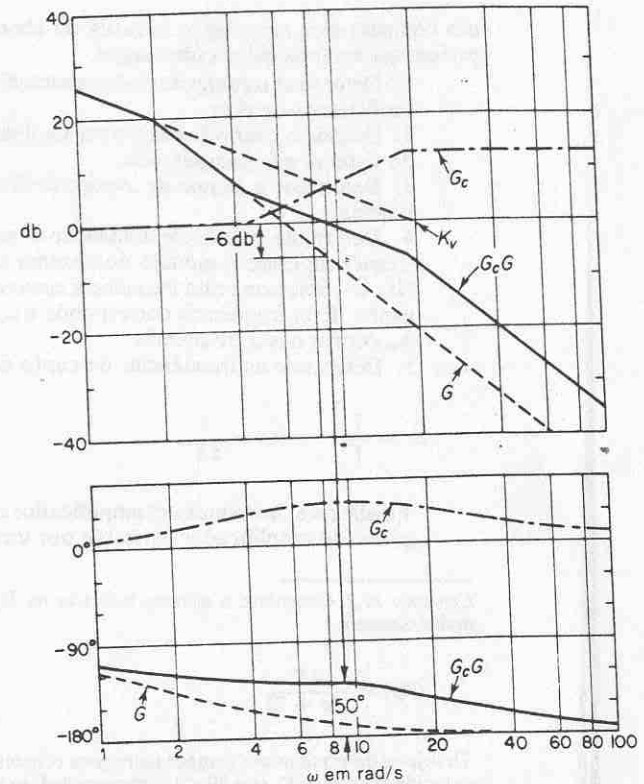


Fig. 10.16 Diagrama de Bode para o sistema compensado.

compensado indicado na Fig. 10.17 satisfaz, portanto, tanto as exigências em regime estacionário como de estabilidade relativa.

Note que para sistemas do tipo 1, como o sistema considerado (com os valores aqui considerados), o valor do coeficiente de erro de velocidade estático K_v é simplesmente o valor da freqüência correspondente à intersecção da reta inicial com inclinação -20 db/década com a reta 0-db, conforme indicado na Fig. 10.16.

A Fig. 10.18 mostra os gráficos polares de $G(j\omega)$ (com $K = 10$) e de $G_c(j\omega)G(j\omega)$. Da Fig. 10.18, verificamos que a freqüência de ressonância do sistema não compensado é aproximadamente 6 rad/s e que a do sistema compensado é aproximadamente 7 rad/s. (Este fato também indica que a largura de faixa foi aumentada.)

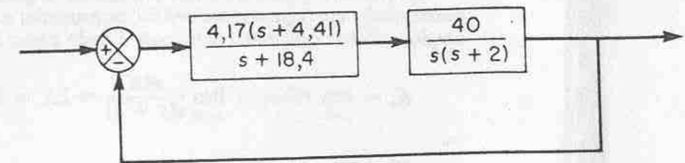


Fig. 10.17 Sistema compensado.

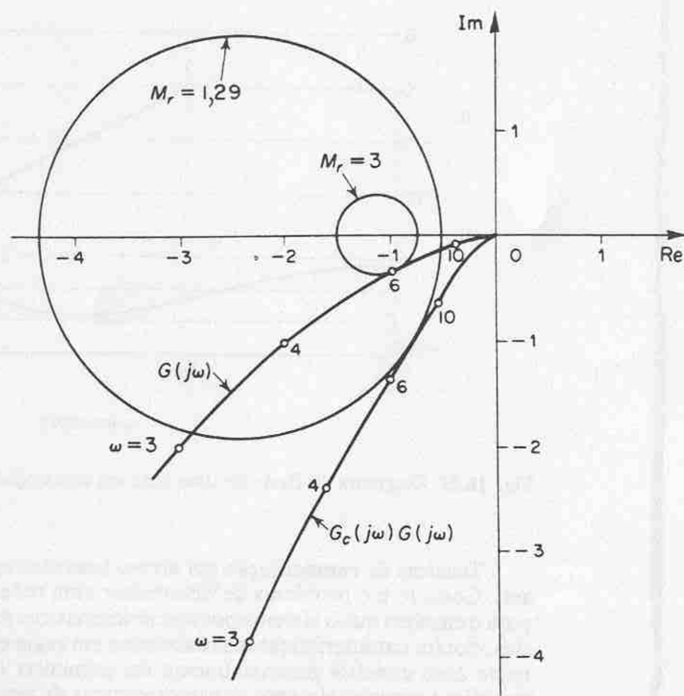


Fig. 10.18 Diagramas polares da função de transferência de malha-aberta não compensada e compensada. (G: sistema não compensado, G_cG: sistema compensado.)

Da Fig. 10.18, verificamos que o valor do pico de ressonância M_r para o sistema não compensado com $K = 10$ é 3. O valor de M_r para o sistema compensado é obtido em 1,29. Isto mostra claramente que o sistema compensado apresentou uma estabilidade relativa melhorada. Note que o valor de M_r pode ser facilmente obtido transferindo-se os dados do diagrama de Bode para a carta de Nichols. Veja Exemplo 10.4.)

Note que se o ângulo de fase de $G(j\omega)$ diminui rapidamente próximo à freqüência de cruzamento de ganho, uma compensação em avanço resulta ineficiente devido ao deslocamento da freqüência de cruzamento de ganho para a direita resultar em uma dificuldade em obter-se suficiente avanço de fase na nova freqüência de cruzamento de ganho. Isto significa que, a fim de propiciar a margem de fase desejada, devemos usar um valor muito pequeno para α. O valor de α, entretanto, não deve ser menor do que 0,07 nem deve propiciar um ângulo de fase em avanço máximo ϕ_m maior do que 60° porque estes valores exigirão um ganho adicional de valor excessivo. (Se for necessário mais do que 60°, duas (ou mais) redes em avanço devem ser utilizadas em série com um amplificador de isolação.)

10.4 COMPENSAÇÃO EM ATRASO

Redes em atraso. A Fig. 10.19(a) mostra uma rede elétrica de atraso. A denominação "rede em atraso" provém do fato que quando a tensão de entrada e_i é

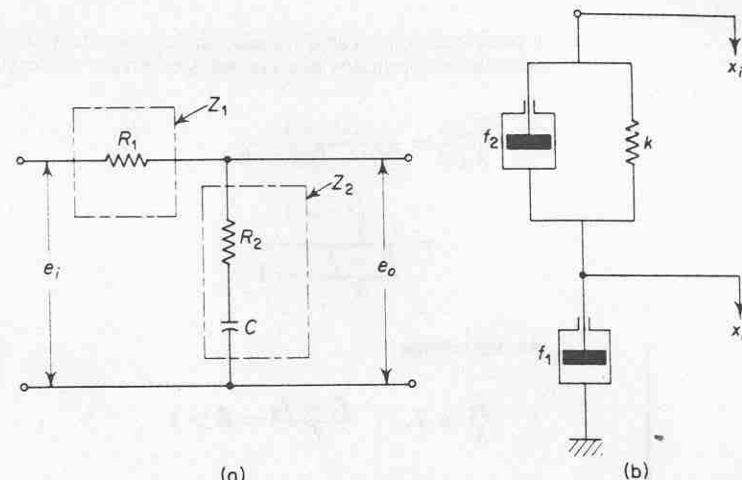


Fig. 10.19 (a) Rede elétrica de atraso; (b) rede mecânica de atraso.

senoidal, a tensão de saída e_0 é senoidal porém atrasada em relação à entrada de um ângulo que é função da freqüência da entrada senoidal. As impedâncias complexas Z_1 e Z_2 são

$$Z_1 = R_1, \qquad Z_2 = R_2 + \frac{1}{Cs}$$

A função de transferência entre a tensão de saída $E_0(s)$ e a tensão de entrada $E_i(s)$ é dada por

$$\frac{E_0(s)}{E_i(s)} = \frac{Z_2}{Z_1 + Z_2} = \frac{R_2Cs + 1}{(R_1 + R_2)Cs + 1}$$

Definindo,

$$R_2C = T, \qquad \frac{R_1 + R_2}{R_2} = \beta > 1$$

Então, a função de transferência resulta em

$$\frac{E_0(s)}{E_i(s)} = \frac{Ts + 1}{\beta Ts + 1} = \frac{1}{\beta}\left(\frac{s + \frac{1}{T}}{s + \frac{1}{\beta T}}\right)$$

A Fig. 10.19(b) mostra uma rede mecânica de atraso. Ela consiste em uma mola e dois amortecedores viscosos. A equação diferencial para esta rede mecânica é

$$f_1\dot{x}_0 = k(x_i - x_0) + f_2(\dot{x}_i - \dot{x}_0)$$

Considerando as transformadas de Laplace de ambos os membros desta equação, admitindo condições iniciais nulas e, então, reescrevendo, obtemos

$$\frac{X_0(s)}{X_i(s)} = \frac{f_2 s + k}{(f_1 + f_2)s + k}$$

$$= \frac{\frac{f_2}{k}s + 1}{\frac{f_1 + f_2}{k}s + 1}$$

Se definirmos

$$\frac{f_2}{k} = T, \qquad \frac{f_1 + f_2}{f_2} = \beta > 1$$

então, a função de transferência $X_0(s)/X_i(s)$ resulta em

$$\frac{X_0(s)}{X_i(s)} = \frac{Ts + 1}{\beta Ts + 1} = \frac{1}{\beta}\left(\frac{s + \frac{1}{T}}{s + \frac{1}{\beta T}}\right)$$

Características de rede de atraso. Uma rede de atraso possui a seguinte função de transferência:

$$\frac{Ts + 1}{\beta Ts + 1} = \frac{1}{\beta}\left(\frac{s + \frac{1}{T}}{s + \frac{1}{\beta T}}\right) \qquad (\beta > 1)$$

No plano complexo, uma rede de atraso possui um pólo em $s = -1/(\beta T)$ e um zero em $s = -1/T$. (O pólo está localizado à direita do zero.)

A Fig. 10.20 mostra um gráfico polar típico de uma rede em atraso. A Fig. 10.21 mostra o diagrama de Bode de uma rede de atraso quando $\beta = 10$. As freqüências de canto da rede de atraso são $\omega = 1/T$ e $\omega = 1/(\beta T)$. Como pode ser verificado na Fig. 10.21, a rede em atraso é essencialmente um filtro passa-baixas.

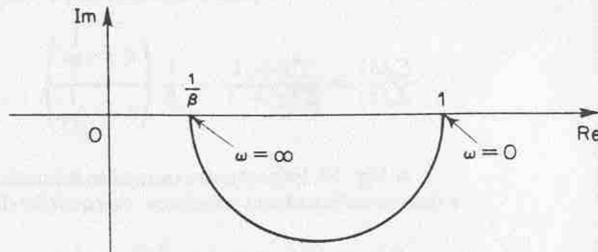


Fig. 10.20 Diagrama polar de uma rede em atraso.

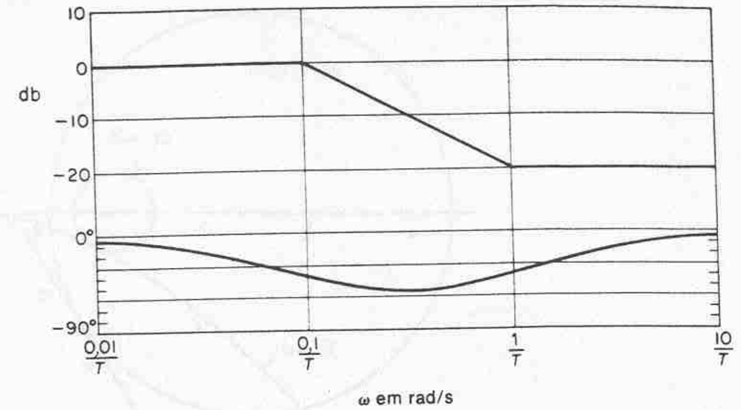


Fig. 10.21 Diagrama de Bode de uma rede em atraso $(j\omega T + 1)/(j\omega\beta T + 1)$ com $\beta = 10$.

Técnicas de compensação em atraso baseadas na abordagem do lugar das raízes. Considere o problema de determinar uma rede de compensação conveniente para o caso em que o sistema possui características de resposta transitória satisfatórias, porém características insatisfatórias em regime permanente. A compensação neste caso consiste essencialmente em aumentar o ganho em malha-aberta sem modificar apreciavelmente as características de resposta transitória. Isto significa que o lugar das raízes na vizinhança dos pólos de malha-fechada dominantes não deve ser modificado significativamente, porém o ganho de malha-aberta deve ser aumentado tanto quanto necessário. Este resultado pode ser conseguido se for colocado um compensador em atraso em cascata com a função de transferência do ramo direto fornecida.

Para evitar uma variação apreciável nos lugares das raízes, a contribuição angular da rede em atraso deve ser limitada a um pequeno valor, digamos 5°. Para assegurar esta hipótese, colocaremos o pólo e o zero da rede em atraso relativamente próximo entre si e próximos da origem do plano s. Então, os pólos de malha-fechada do sistema compensado serão deslocados apenas ligeiramente de suas posições originais. Conseqüentemente, as características de resposta transitória permanecerão essencialmente as mesmas.

Note que se colocarmos o pólo e o zero da rede em atraso muito próximos entre si, então $s_1 + (1/T)$ e $s_1 + (1/\beta T)$ serão quase iguais, onde s_1 é o pólo de malha-fechada. Portanto,

$$\left|\frac{1}{\beta}\left(\frac{s_1 + \frac{1}{T}}{s_1 + \frac{1}{\beta T}}\right)\right| \doteq \frac{1}{\beta}$$

Isto implica que o ganho de malha-aberta pode ser aumentado aproximadamente por um fator de β sem alteração nas características de resposta transitória. Se o pólo e o zero forem colocados muito próximos à origem, o valor de β pode ser feito grande. Normalmente, $1 < \beta < 15$, e $\beta = 10$ é uma boa escolha.

Um aumento no ganho significa um aumento nos coeficientes de erro estático. Se a função de transferência de malha-aberta do sistema não compensado é $G(s)$, então o coeficiente de erro de velocidade estático K_v é

$$K_v = \lim_{s\to 0} sG(s)$$

Se o compensador for escolhido como

$$G_c(s) = K_c \frac{Ts + 1}{\beta Ts + 1} = \frac{K_c}{\beta}\left(\frac{s + \frac{1}{T}}{s + \frac{1}{\beta T}}\right) \tag{10.3}$$

então, para o sistema compensado com a função de transferência de malha-aberta $G_c(s)G(s)$, o coeficiente de erro de velocidade estático $\hat{K}_v$ resulta em

$$\hat{K}_v = \lim_{s\to 0} sG_c(s)G(s) = \lim_{s\to 0} G_c(s)K_v = K_c K_v$$

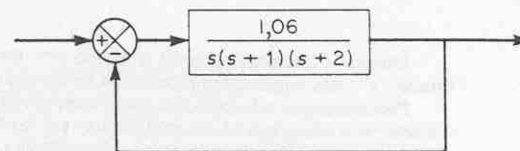
Portanto, se o compensador for dado pela Eq. (10.3), então o coeficiente de erro de velocidade estático é aumentado por um fator de K_c.

Os procedimentos para projeto de um compensador em atraso pelo método do lugar das raízes podem ser estabelecidos como segue: (Suporemos que o sistema não compensado satisfaz as especificações de resposta transitória através de um simples ajuste do ganho. Se este não for o caso, referir-se à Seção 10.5.)

1. Desenhe o gráfico do lugar das raízes para o sistema não compensado. Com base nas especificações da resposta transitória, localize os pólos de malha-fechada dominantes sobre o lugar das raízes.
2. Determine o ganho de malha-aberta utilizando a condição do módulo.
3. Calcule o particular coeficiente de erro especificado no problema.
4. Determine a quantidade de aumento no coeficiente de erro necessária para satisfazer as especificações.
5. Determine o pólo e o zero da rede em atraso que produz o aumento necessário no coeficiente particular de erro sem alteração apreciável nos lugares das raízes originais.
6. Desenhe um novo gráfico do lugar das raízes para o sistema compensado. Localize os pólos de malha-fechada dominantes sobre o lugar das raízes. (Se a contribuição angular da rede em atraso for muito pequena, isto é, poucos graus, então os lugares das raízes original e modificado são praticamente idênticos. Haverá uma leve discrepância entre eles. Localize, então, sobre o novo lugar das raízes os pólos de malha-fechada dominantes desejados com base nas especificações de resposta transitória.)
7. Ajuste o ganho do amplificador por meio da condição do módulo para que os pólos de malha-fechada dominantes resultem nas posições desejadas.

Exemplo 10.3 Considere o sistema indicado na Fig. 10.22(a). A função de transferência do ramo direto é

$$G(s) = \frac{1{,}06}{s(s+1)(s+2)}$$



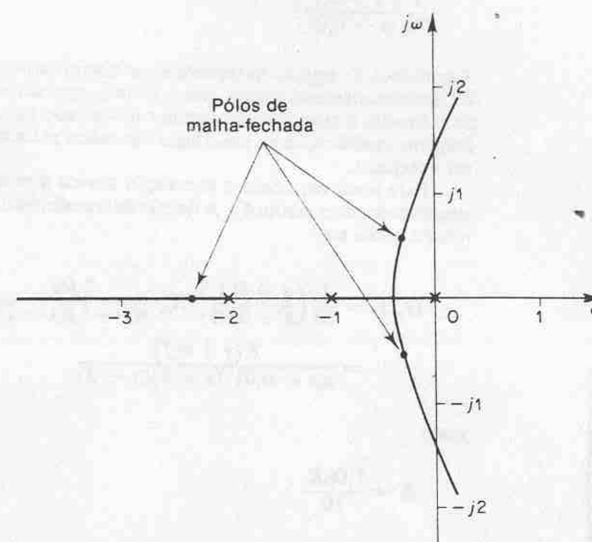


Fig. 10.22 (a) Sistema de controle; (b) gráfico do lugar das raízes.

O gráfico do lugar das raízes para o sistema é fornecido na Fig. 10.22(b). A função de transferência de malha-fechada resulta em

$$\frac{C(s)}{R(s)} = \frac{1{,}06}{s(s+1)(s+2) + 1{,}06}$$
$$= \frac{1.06}{(s + 0{,}33 - j0{,}58)(s + 0{,}33 + j0{,}58)(s + 2{,}33)}$$

Os pólos de malha-fechada dominantes são

$$s = -0{,}33 \pm j0{,}58$$

A relação de amortecimento dos pólos de malha-fechada dominantes é $\zeta = 0{,}5$. A freqüência natural não amortecida dos pólos de malha-fechada dominantes é 0,67 rad/s. O coeficiente de erro de velocidade estático é 0,53 s^{-1}.

Deseja-se aumentar o coeficiente de erro de velocidade estático K_v para aproximadamente 5 s^{-1} sem mudança apreciável na localização dos pólos de malha-fechada dominantes.

Para satisfazer esta especificação, vamos inserir um compensador em atraso, que consiste em uma rede em atraso e um amplificador, em cascata com a função de transferência do ramo direto fornecida. Para aumentar o coeficiente de erro de velocidade estático por um fator de aproximadamente 10, vamos colocar o pólo e o zero da rede em atraso em $s \approx -0{,}01$ e $s = -0{,}1$, respectivamente. A função de transferência da rede em atraso é então

$$\frac{1}{10}\left(\frac{s+0{,}1}{s+0{,}01}\right)$$

A contribuição angular desta rede em atraso próxima a um pólo de malha-fechada dominante é de aproximadamente alguns graus. (Isto é, aproximadamente o máximo que podemos permitir.) Devido à contribuição angular desta rede em atraso não ser muito pequena, há uma pequena modificação no novo lugar das raízes próximo aos pólos de malha-fechada dominantes desejados.

Para levar em conta a atenuação devida à rede em atraso, colocamos em cascata um amplificador com ganho K_c. A função de transferência do ramo direto do sistema compensado resulta então em

$$G_1(s) = \frac{1}{10}\left(\frac{s+0{,}1}{s+0{,}01}\right)(K_c)\frac{1{,}06}{s(s+1)(s+2)}$$

$$= \frac{K(s+0{,}1)}{s(s+0{,}01)(s+1)(s+2)}$$

onde

$$K = \frac{1{,}06K_c}{10}$$

O diagrama de blocos do sistema compensado é indicado na Fig. 10.23(a). O gráfico do lugar das raízes para o sistema compensado próximo aos pólos de malha-fechada dominantes é indicado na Fig. 10.23(b), conjuntamente com o lugar das raízes original.

Se a relação de amortecimento dos novos pólos de malha-fechada dominantes for mantida a mesma, então os pólos são obtidos do novo gráfico do lugar das raízes como segue:

$$s_1 = -0{,}28 + j0{,}51,$$

$$s_2 = -0{,}28 - j0{,}51$$

O ganho de malha-aberta K é

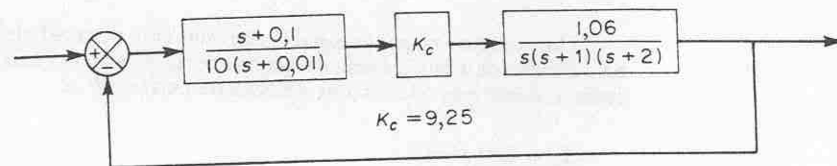
$$K = \left|\frac{s(s+0{,}01)(s+1)(s+2)}{s+0{,}1}\right|_{s=-0{,}28+j0{,}51} = 0{,}98$$

Portanto, o ganho do amplificador K_c é

$$K_c = \frac{10}{1{,}06}K = 9{,}25$$

Então, o sistema compensado possui a seguinte função de transferência de malha-aberta:

$$G_1(s) = \frac{0{,}98(s+0{,}1)}{s(s+0{,}01)(s+1)(s+2)} = \frac{4{,}9(10s+1)}{s(100s+1)(s+1)(0{,}5s+1)}$$



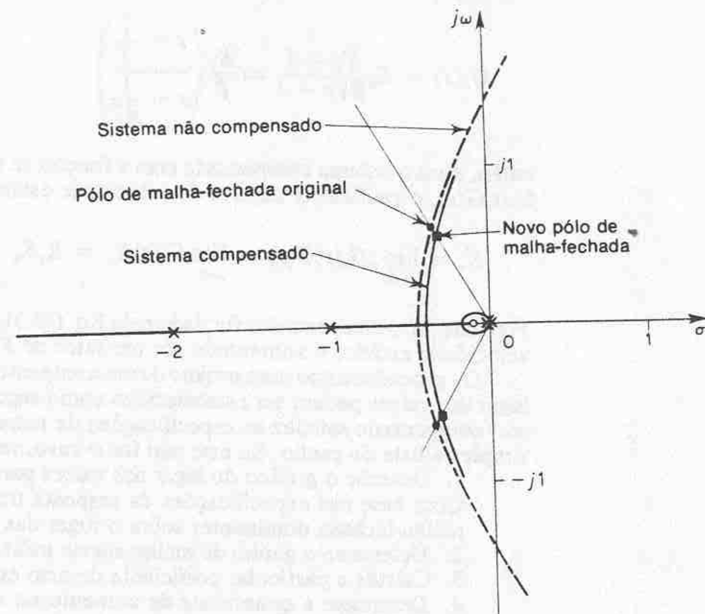


Fig. 10.23 (a) Sistema compensado; (b) gráficos dos lugares das raízes para o sistema compensado e para o sistema não compensado.

O coeficiente de erro de velocidade estático K_v é

$$K_v = \lim_{s\to 0} sG_1(s) = 4{,}9\ s^{-1}$$

No sistema compensado, o coeficiente de erro de velocidade estático foi aumentado para 4,9 s^{-1}, ou 4,9/0,53 = 9,25 vezes o valor original. (O erro de regime estacionário com entradas em rampa foi diminuído de aproximadamente 11% daquele existente no sistema original.) Conseguimos essencialmente atingir o objetivo do projeto de aumentar o coeficiente de erro de velocidade estático para aproximadamente 5 s^{-1}. (Se desejarmos aumentar o coeficiente de erro de velocidade estático para exatamente 5 s^{-1}, ou podemos modificar as posições do pólo e do zero da rede em atraso ou podemos usar a rede em atraso presente e escolher $K_c = 9{,}44$. Neste último caso, entretanto, a relação de amortecimento dos pólos de malha-fechada dominante é menor do que 0,5. Neste problema, podemos considerar o projeto apresentado como bem aceitável.)

Os dois pólos de malha-fechada restantes do sistema compensado são determinados como segue:

$$s_3 = -2{,}31, \qquad s_4 = -0{,}137$$

A inclusão da rede em atraso aumenta a ordem do sistema de três para quatro, acrescentando-se um pólo adicional próximo ao zero da rede em atraso. Como o pólo adicional em $s = -0{,}137$ está próximo ao zero em $s = -0{,}1$, o efeito deste pólo na resposta transitória é pequeno. Desde que o pólo em $s = -2{,}31$ está muito longe do eixo $j\omega$ comparado com os pólos de malha-fechada dominantes, o efeito deste pólo na resposta transitória também é pequeno. Podemos portanto ignorar, com pequeno erro, os pólos de malha-fechada s_3 e s_4. Conclui-se que os dois pólos de malha-fechada s_1 e s_2 são verdadeiramente dominantes. Podemos prever uma resposta bastante precisa considerando-se apenas os pólos de malha-fechada dominantes.

A freqüência natural não amortecida do sistema compensado é 0,6 rad/s. Este valor é aproximadamente 10% menor do que o valor original, 0,67 rad/s. Isto implica que a resposta transitória do sistema compensado é mais lenta do que a do sistema original. Resultará um maior tempo para acomodação. Se este fato puder ser tolerado, a compensação em atraso, conforme aqui discutida, apresenta uma solução satisfatória para o problema de projeto especificado.

Finalmente, note que os pontos de cruzamento dos lugares das raízes com o eixo $j\omega$ tanto do sistema original como do sistema compensado são aproximadamente os mesmos. Isto significa que o sistema compensado ainda é estável mesmo que o ganho de malha-aberta seja aumentado por um fator de aproximadamente 10 sobre o valor crítico original.

Técnicas de compensação em atraso baseadas na abordagem de resposta em freqüência. A função principal de uma rede em atraso é fornecer atenuação na faixa de alta-freqüência de modo a propiciar uma margem de fase suficiente ao sistema. A característica de atraso de fase não apresenta conseqüências na compensação em atraso.

Os procedimentos para projetar um compensador em atraso, através da abordagem de resposta em freqüência, podem ser estabelecidos como segue: (Suporemos que o sistema possui realimentação unitária.)

1. Determine o ganho de malha-aberta tal que seja satisfeita a exigência do particular coeficiente de erro.
2. Usando o ganho determinado anteriormente, desenhe o diagrama de Bode do sistema não compensado e determine as margens de ganho e de fase deste sistema.
3. Se as especificações relativas às margens de fase e de ganho não estiverem satisfeitas, então, determine o ponto de freqüência onde o ângulo de fase da função de transferência de malha-aberta é igual a $-180°$ mais a margem de fase exigida. A margem de fase exigida é a margem de fase especificada mais 5° a 12°. (A adição de 5° a 12° compensa o atraso de fase da rede em atraso.) Escolha esta freqüência como a nova freqüência de cruzamento do ganho.
4. Escolha a freqüência de canto $\omega = 1/T$ (correspondente ao zero da rede em atraso) uma oitava a uma década abaixo da nova freqüência de cruzamento do ganho. (Se as constantes de tempo da rede em atraso não se tornarem demasiadamente grandes, a freqüência de canto $\omega = 1/T$ pode ser escolhida uma década abaixo da nova freqüência de cruzamento do ganho.)
5. Determine a atenuação necessária para fazer com que a curva de módulo seja tal que o módulo é 0 db para a nova freqüência de cruzamento do ganho. Verificando que esta atenuação é $-20 \log \beta$, determine o valor de β. Então a outra freqüência de canto (correspondente ao pólo da rede em atraso) é determinada a partir de $\omega = 1/(\beta T)$.

Exemplo 10.4 Considere o sistema indicado na Fig. 10.24. A função de transferência de malha-aberta é dada por

$$G(s) = \frac{K}{s(s+1)(0{,}5s+1)}$$

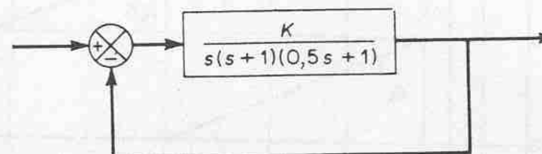


Fig. 10.24 Sistema de controle.

Deseja-se compensar o sistema de modo que o coeficiente de erro de velocidade K_v seja 5 s^{-1}, a margem de fase pelo menos 40°, e a margem de ganho pelo menos 10 db.

O primeiro passo no projeto é ajustar o ganho K para satisfazer o coeficiente de erro de velocidade estático exigido. Portanto,

$$K_v = \lim_{s\to 0} sG(s) = \lim_{s\to 0} \frac{sK}{s(s+1)(0{,}5s+1)} = K = 5$$

ou

$$K = 5$$

Com $K = 5$, o sistema da Fig. 10.24 satisfaz os requisitos de desempenho em regime estacionário.

Construiremos o novo gráfico do diagrama de Bode a partir de

$$G(j\omega) = \frac{5}{j\omega(j\omega+1)(0{,}5j\omega+1)}$$

A curva de módulo e a curva do ângulo de fase de $G(j\omega)$ são indicadas na Fig. 10.25. Deste gráfico, a margem de fase obtida é $-20°$, significando que o sistema é instável.

A compensação em avanço é ineficiente neste caso. Demonstraremos que o uso da compensação em atraso resultará na obtenção das especificações exigidas. Notando que a adição de uma rede em atraso modifica a curva de fase do diagrama de Bode, devemos permitir 5° a 12° para a margem de fase especificada, a fim de compensar a modificação na curva de fase. Desde que a freqüência correspondente a uma margem de fase de 40° é 0,7 rad/s, a nova freqüência de cruzamento do ganho (do sistema compensado) deve ser escolhida próximo a este valor. Para evitar constantes de tempo excessivas para a rede em atraso, escolheremos a freqüência de canto $\omega = 1/T$ (que corresponde ao zero da rede em atraso) como 0,1 rad/s. Desde que esta freqüência de canto não está muito longe abaixo da freqüência de cruzamento do ganho, a modificação na curva de fase não deve ser pequena. Portanto, adicionamos aproximadamente 12° à margem de fase dada para levar em conta o ângulo de atraso introduzido pela rede em atraso. A margem de fase exigida é agora 52°. O ângulo de fase da função de transferência de malha-aberta do sistema não compensado é $-128°$ em aproximadamente $\omega = 0{,}5$ rad/s. Assim, escolhemos a nova freqüência de cruzamento do ganho em 0,5 rad/s. Para que a curva de módulo passe por 0 db nesta nova freqüência de cruzamento do ganho, a rede em atraso deve fornecer a atenuação necessária, que é, neste caso, -20 db.

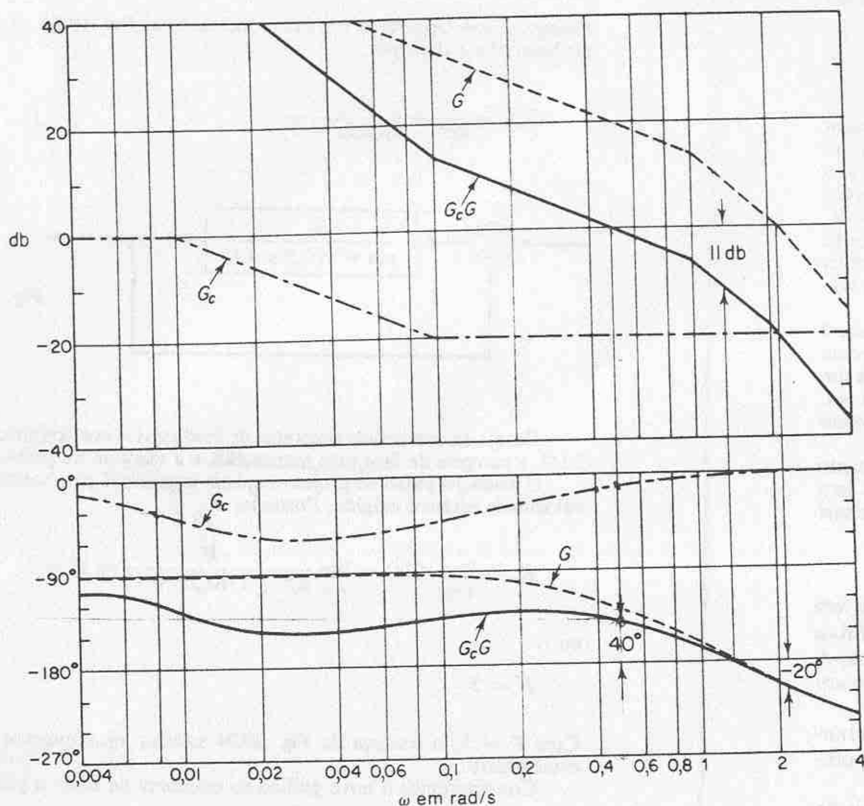


Fig. 10.25 Diagramas de Bode para o sistema não compensado, para o compensador e para o sistema compensado. (G: sistema não compensado. G_c: compensador. G_cG: sistema compensado.)

Portanto,

$$20 \log \frac{1}{\beta} = -20$$

ou

$$\beta = 10$$

A outra freqüência de canto $\omega = 1(\beta T)$ (que corresponde ao pólo da rede em atraso) é então determinada como

$$\frac{1}{\beta T} = 0{,}01 \text{ rad/s}$$

Portanto, a função de transferência da rede em atraso necessária é dada por

$$\frac{1}{10}\frac{s + 0{,}1}{s + 0{,}01}$$

A função de transferência de malha-aberta do sistema compensado é

$$G_c(s)G(s) = \frac{5(10s + 1)}{s(100s + 1)(s + 1)(0{,}5s + 1)}$$

As curvas de módulo e de ângulo de fase de $G_c(j\omega)G(j\omega)$ também são indicadas na Fig. 10.25. Note que o efeito no atraso de fase da rede é desprezível em altas freqüências.

A margem de fase do sistema compensado é aproximadamente 40°, que é o valor exigido. A margem de ganho é aproximadamente 11 db, que é bem aceitável. O coeficiente de erro de velocidade estático é $5\ s^{-1}$ conforme exigido. O sistema compensado, portanto, satisfaz as especificações tanto em relação ao regime estacionário como de estabilidade relativa.

Note que a nova freqüência de cruzamento do ganho foi diminuída de 2,1 para 0,5 rad/s. Isto significa que a largura de faixa do sistema foi reduzida. Portanto, a velocidade da resposta transitória do sistema compensado será mais lenta do que aquela do sistema original.

Para mostrar os efeitos da compensação em atraso, os gráficos do log-módulo *versus* fase dos sistemas não compensado (com $K = 5$) e compensado são indicados na Fig. 10.26. O gráfico de $G(j\omega)$ mostra claramente que o sistema não compensado é instável. A adição da rede em atraso estabiliza o sistema. O gráfico de $G_c(j\omega)G(j\omega)$ é tangente ao lugar geométrico de $M = 3$ db. Portanto, o valor do pico de ressonância M_r é 3 db, ou 1,4, e este pico ocorre em $\omega = 0{,}5$ rad/s.

Comparando o sistema compensado aqui projetado com aquele projetado no Exemplo 10.3, notamos que são muito parecidos. Este resultado, entretanto, pode não ser necessariamente verdadeiro em outros casos. Compensadores projetados por métodos diferentes ou por

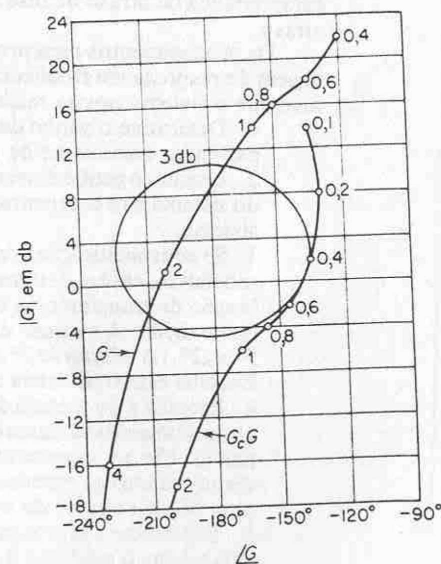


Fig. 10.26 Gráficos do log-módulo *versus* fase do sistema não compensado e do sistema compensado. (G: sistema não compensado, G_cG: sistema compensado.)

projetistas diferentes (ainda que utilizando a mesma abordagem) podem resultar significativamente diferentes. Qualquer sistema bem projetado, entretanto, resultará em especificações em regime estacionário e regime transitório similares. A melhor entre muitas alternativas pode ser escolhida a partir da consideração econômica de que as constantes de tempo da rede em atraso não devem ser muito grandes desde que os valores e dimensões dos resistores exigidos são relacionados diretamente com os valores das constantes de tempo.

Alguns poucos comentários sobre a compensação em atraso

1. As redes em atraso são essencialmente filtros passa-baixas. Portanto, a compensação em atraso permite um alto ganho em baixas freqüências (o que melhora o erro em regime permanente) e reduz o ganho na faixa crítica de maiores freqüências de modo a evitar a instabilidade do sistema. Note que na compensação em atraso, utilizamos a característica de atenuação da rede em atraso em altas freqüências ao invés da característica de atraso de fase. (A característica de atraso de fase não é utilizado para fins de compensação.)
2. A atenuação devida à rede em atraso desloca a freqüência de cruzamento do ganho para um ponto de freqüência menor, onde a margem de fase é aceitável. Portanto, a rede em atraso reduz a largura de faixa do sistema resultando em uma resposta transitória mais lenta. [A curva do ângulo de fase de $G_c(j\omega)G(j\omega)$ é relativamente inalterada próximo e acima da nova freqüência de cruzamento de ganho.]
3. Desde que um compensador em atraso tende a integrar o sinal de entrada, ele age aproximadamente como um controlador proporcional-mais-integral.

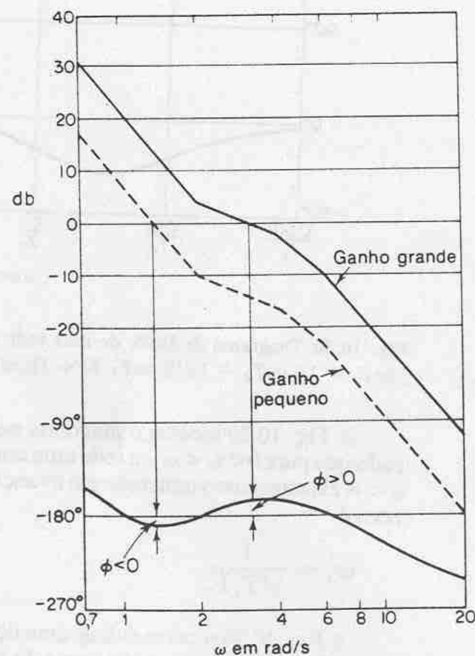


Fig. 10.27 Diagrama de Bode de um sistema condicionalmente estável.

Por esta razão, um sistema compensado em atraso tende a tornar-se menos estável. Para evitar esta característica indesejável, a constante de tempo T deve ser suficientemente maior do que a maior constante de tempo do sistema.

4. Pode ocorrer estabilidade condicional quando um sistema, possuindo saturação ou limitação, é compensado através de um compensador em atraso. Quando ocorrer a saturação ou limitação no sistema, o compensador reduz o ganho de malha efetivo. Então, o sistema torna-se menos estável, podendo resultar uma operação instável, como indicada na Fig. 10.27. Para evitar esta ocorrência, o sistema deve ser projetado de modo que o efeito da compensação em atraso se torne significativo somente quando a amplitude da entrada para o elemento saturante for pequena. (Isto pode ser feito por meio de uma compensação por laço de realimentação interna.)

10.5 COMPENSAÇÃO AVANÇO-ATRASO

A compensação em avanço aumenta a largura de faixa, o que melhora a velocidade de resposta, e também reduz a sobrelevação. Entretanto, a melhora no desempenho de regime permanente é muito pequena. A compensação em atraso resulta em uma grande melhoria no desempenho de regime permanente porém resulta em uma resposta mais lenta, devido à largura de faixa reduzida.

Se forem desejadas melhorias tanto na resposta transitória como na resposta em regime estacionário (isto é, um aumento significativo no ganho e na largura de faixa), então devem ser utilizadas simultaneamente uma rede em avanço e uma rede em atraso. Ao invés de introduzir uma rede em avanço e uma rede em atraso como elementos separados, entretanto, é mais econômico utilizar uma única rede avanço-atraso. A rede avanço-atraso combina as vantagens das redes em avanço e em atraso.

A rede em avanço-atraso possui dois pólos e dois zeros. Portanto, esta compensação aumenta a ordem do sistema por duas unidades, a menos que ocorra cancelamento de um pólo e um zero no sistema compensado.

Redes avanço-atraso. A Fig. 10.28 mostra uma rede elétrica de avanço-atraso. Para uma entrada senoidal a saída é senoidal com uma defasagem que é função da freqüência de entrada. Este ângulo de fase varia de atrasado para avançado conforme a freqüência é aumentada desde zero a infinito. Note que o avanço e o atraso de fase ocorrem em diferentes faixas de freqüência.

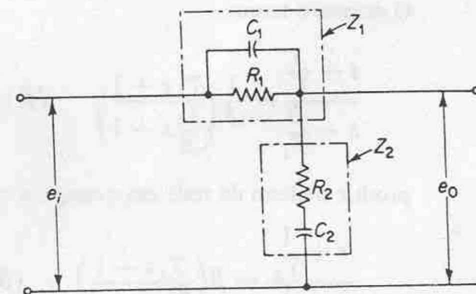


Fig. 10.28 Rede elétrica avanço-atraso.

Obteremos a seguir a função de transferência da rede avanço-atraso. As impedâncias complexas Z_1 e Z_2 são

$$Z_1 = \frac{R_1}{R_1C_1s+1}, \qquad Z_2 = R_2 + \frac{1}{C_2s}$$

A função de transferência entre $E_0(s)$ e $E_i(s)$ é

$$\frac{E_0(s)}{E_i(s)} = \frac{Z_2}{Z_1+Z_2} = \frac{(R_1C_1s+1)(R_2C_2s+1)}{(R_1C_1s+1)(R_2C_2s+1)+R_1C_2s}$$

O denominador desta função de transferência pode ser fatorado em dois termos reais. Vamos definir

$$R_1C_1 = T_1, \qquad R_2C_2 = T_2$$

$$R_1C_1 + R_2C_2 + R_1C_2 = \frac{T_1}{\beta} + \beta T_2 \qquad (\beta > 1)$$

Então, $E_0(s)/E_i(s)$ pode ser simplificada para

$$\frac{E_0(s)}{E_i(s)} = \frac{(T_1s+1)(T_2s+1)}{\left(\frac{T_1}{\beta}s+1\right)(\beta T_2s+1)} = \frac{\left(s+\frac{1}{T_1}\right)\left(s+\frac{1}{T_2}\right)}{\left(s+\frac{\beta}{T_1}\right)\left(s+\frac{1}{\beta T_2}\right)}$$

Características da rede avanço-atraso. Considere a função de transferência da rede avanço-atraso

$$\left(\frac{s+\frac{1}{T_1}}{s+\frac{\beta}{T_1}}\right)\left(\frac{s+\frac{1}{T_2}}{s+\frac{1}{\beta T_2}}\right)$$

O primeiro termo

$$\frac{s+\frac{1}{T_1}}{s+\frac{\beta}{T_1}} = \frac{1}{\beta}\left(\frac{T_1s+1}{\frac{T_1}{\beta}s+1}\right) \qquad (\beta > 1)$$

produz o efeito da rede em avanço, e o segundo termo

$$\frac{s+\frac{1}{T_2}}{s+\frac{1}{\beta T_2}} = \beta\left(\frac{T_2s+1}{\beta T_2s+1}\right) \qquad (\beta > 1)$$

produz o efeito da rede em atraso.

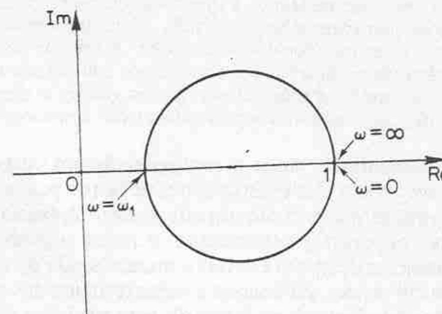


Fig. 10.29 Diagrama polar de uma rede avanço-atraso.

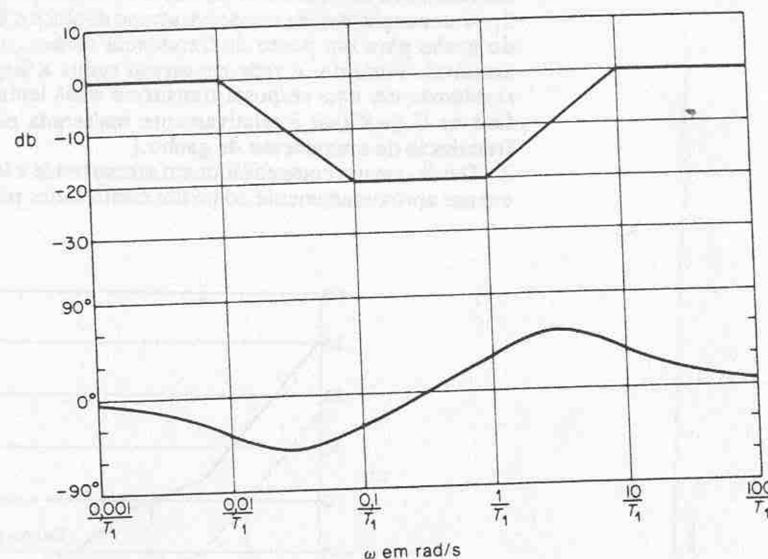


Fig. 10.30 Diagrama de Bode de uma rede avanço-atraso. $(j\omega T_1 + 1)(j\omega T_2 + 1)/\{[(j\omega T_1/\beta) + 1](j\omega\beta T_2 + 1)\}$ com $\beta = 10$ e $T_2 = 10T_1$.

A Fig. 10.29 mostra o diagrama polar da rede avanço-atraso. Pode ser verificado que para $0 < \omega < \omega_1$, a rede atua como uma rede em atraso, enquanto para $\omega_1 < \omega < \infty$ ela atua como uma rede em avanço. A freqüência ω_1 na qual o ângulo de fase é zero é

$$\omega_1 = \frac{1}{\sqrt{T_1T_2}}$$

A Fig. 10.30 mostra o diagrama de Bode de uma rede avanço-atraso quando $\beta = 10$ e $T_2 = 10T_1$. Note que a curva de módulo apresenta o valor 0 db nas regiões de baixa e de alta freqüências, em virtude de a função de transferência da rede avanço-atraso como um todo não conter β como fator.

Técnicas de compensação de avanço-atraso baseadas na abordagem do lugar das raízes. Os procedimentos para projetar um compensador avanço-atraso podem ser estabelecidos como segue:

1. A partir das especificações de desempenho fornecidas, determine a localização desejada para os pólos de malha-fechada dominantes.
2. Para obter os pólos de malha-fechada dominantes nas posições desejadas, calcule a contribuição do ângulo ϕ necessária para a parte do avanço de fase da rede avanço-atraso.
3. Usando a seguinte função de transferência do compensador avanço-atraso:

$$G_c(s) = \left(\frac{s + \frac{1}{T_1}}{s + \frac{\beta}{T_1}}\right)\left(\frac{s + \frac{1}{T_2}}{s + \frac{1}{\beta T_2}}\right)K_c$$

determine a constante K_c a partir das exigências do coeficiente de erro particular especificado no problema de projeto.

4. Para o compensador avanço-atraso, escolha T_2 suficientemente grande de modo que

$$\left|\frac{s_1 + \frac{1}{T_2}}{s_1 + \frac{1}{\beta T_2}}\right|$$

seja aproximadamente unitário, onde $s = s_1$ é um dos pólos de malha-fechada dominante. Determine os valores de T_1 e β das exigências que

$$\left|\frac{s_1 + \frac{1}{T_1}}{s_1 + \frac{\beta}{T_1}}\right| |K_c G(s_1)| = 1$$

$$\angle \frac{s + \frac{1}{T_1}}{s + \frac{\beta}{T_1}} = \phi$$

5. Usando o valor de β determinado anteriormente, escolha T_2 de modo que

$$\left|\frac{s_1 + \frac{1}{T_2}}{s_1 + \frac{1}{\beta T_2}}\right| \doteq 1 \qquad 0 < \angle \frac{s_1 + \frac{1}{T_2}}{s_1 + \frac{1}{\beta T_2}} < 3°$$

O valor de βT_2, a maior constante de tempo da rede avanço-atraso, não deve ser demasiadamente grande para ser fisicamente realizável.

Exemplo 10.5 Considere o sistema de controle indicado na Fig. 10.31. A função de transferência do ramo direto é

$$G(s) = \frac{4}{s(s + 0{,}5)}$$

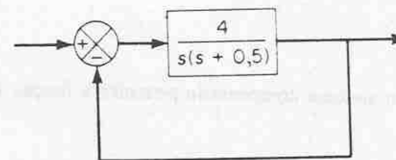


Fig. 10.31 Sistema de controle.

Este sistema possui pólos de malha-fechada em

$$s = -0{,}25 \pm j1{,}98$$

A relação de amortecimento é 0,125, a freqüência natural não amortecida é 2 rad/s, e o coeficiente de erro de velocidade estático é 8 s^{-1}.

Deseja-se que a relação de amortecimento dos pólos de malha-fechada dominantes seja igual a 0,5 e também que aumente a freqüência natural não amortecida para 5 rad/s e o coeficiente de erro de velocidade estático para 50 s^{-1}. Projete um compensador apropriado para satisfazer todas as especificações de desempenho.

A partir das especificações de desempenho, os pólos de malha-fechada dominantes devem ser

$$s = -2{,}50 \pm j4{,}33$$

Desde que,

$$\angle \frac{4}{s(s + 0{,}5)}\bigg|_{s=-2{,}50+j4{,}33} = -235°$$

a parte do avanço de fase da rede avanço-atraso deve contribuir com 55° de modo que o lugar das raízes passe através das posições desejadas dos pólos de malha-fechada dominantes.

O compensador avanço-atraso possui a função de transferência

$$G_c(s) = \left(\frac{s + \frac{1}{T_1}}{s + \frac{\beta}{T_1}}\right)\left(\frac{s + \frac{1}{T_2}}{s + \frac{1}{\beta T_2}}\right)K_c$$

Portanto o sistema compensado terá a função de transferência

$$G_c(s)G(s) = \left(\frac{s + \frac{1}{T_1}}{s + \frac{\beta}{T_1}}\right)\left(\frac{s + \frac{1}{T_2}}{s + \frac{1}{\beta T_2}}\right)K_c G(s)$$

Desta forma, o coeficiente de erro de velocidade estático resulta em

$$K_v = \lim_{s\to 0} sG_c(s)G(s) = \lim_{s\to 0} sK_cG(s)$$

O requisito sobre o coeficiente de erro de velocidade estático é $K_v = 50\ \mathrm{s}^{-1}$. Portanto,

$$K_v = \lim_{s\to 0} sK_cG(s) = \lim_{s\to 0} \frac{s4K_c}{s(s+0{,}5)} = 8K_c = 50$$

Portanto, K_c resulta

$$K_c = 6{,}25$$

Conseqüentemente, o sistema compensado possuirá a função de transferência de malha-aberta

$$G_c(s)G(s) = \left(\frac{s + \dfrac{1}{T_1}}{s + \dfrac{\beta}{T_1}}\right)\left(\frac{s + \dfrac{1}{T_2}}{s + \dfrac{1}{\beta T_2}}\right)\frac{25}{s(s+0{,}5)}$$

Desde que escolhemos T_2 suficientemente grande para

$$\left|\frac{s + \dfrac{1}{T_2}}{s + \dfrac{1}{\beta T_2}}\right|_{s=-2,5+j4,33} \doteqdot 1$$

se desejarmos que os pólos de malha-fechada estejam em $s = -2.5 \pm j4.33$, a condição do módulo resulta em

$$|G_c(s)G(s)|_{s=-2,5+j4,33} = \left|\frac{s + \dfrac{1}{T_1}}{s + \dfrac{\beta}{T_1}}\right|\left|\frac{25}{s(s+0{,}5)}\right|_{s=-2,5+j4,33} = \left|\frac{s + \dfrac{1}{T_1}}{s + \dfrac{\beta}{T_1}}\right|\frac{5}{4{,}77} = 1$$

e a condição do ângulo resulta em

$$\angle\left.\frac{s + \dfrac{1}{T_1}}{s + \dfrac{\beta}{T_1}}\right|_{s=-2,5+j4,33} = 55°$$

É relativamente fácil determinar graficamente os valores de T_1 e β que satisfazem estas condições de módulo e ângulo. Referindo-se à Fig. 10.32, podemos facilmente localizar os pontos A e B de modo que

$$\angle APB = 55°, \qquad \frac{\overline{PA}}{\overline{PB}} = \frac{4{,}77}{5}$$

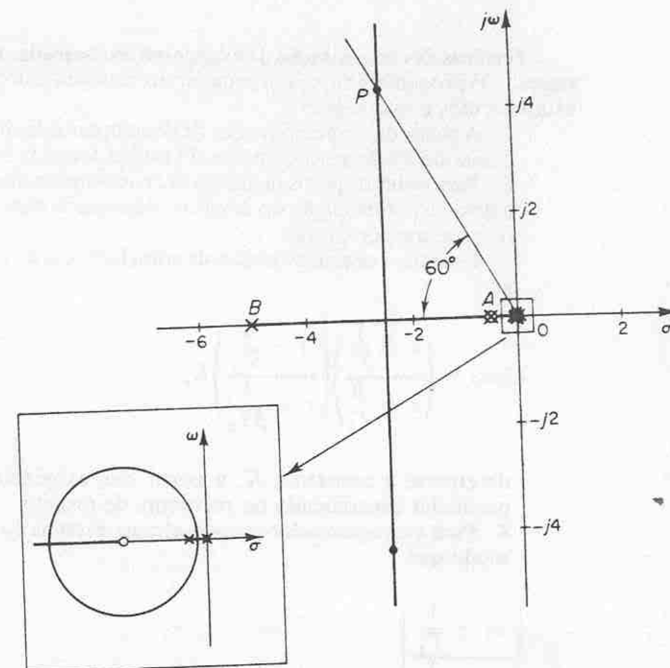


Fig. 10.32 Gráfico do lugar das raízes do sistema compensado.

Graficamente, a partir da Fig. 10.32, obtemos

$$\overline{AO} = 0{,}5, \qquad \overline{BO} = 5$$

Então,

$$-\frac{1}{T_1} = -0{,}5, \qquad -\frac{\beta}{T_1} = -5$$

Portanto,

$$T_1 = 2, \qquad \beta = 10$$

Conseqüentemente, a parte do avanço de fase da rede avanço-atraso resulta em

$$\frac{s+0{,}5}{s+5}$$

Para a parte do atraso de fase da rede avanço-atraso, é exigido que

$$\left|\frac{s + \dfrac{1}{T_2}}{s + \dfrac{1}{10T_2}}\right|_{s=-2,5+j4,33} \doteqdot 1 \qquad 0 < \angle\left.\frac{s + \dfrac{1}{T_2}}{s + \dfrac{1}{10T_2}}\right|_{s=-2,5+j4,33} < 3°$$

Para satisfazer estas relações e, ao mesmo tempo, assegurar que a maior constante de tempo ($10T_2$) da rede avanço-atraso não seja demasiadamente grande para ser fisicamente realizável, escolhemos

$$T_2 = 10$$

Então, a função de transferência do compensador avanço-atraso é

$$G_c(s) = \left(\frac{s+0,5}{s+5}\right)\left(\frac{s+0,1}{s+0,01}\right)(6,25)$$



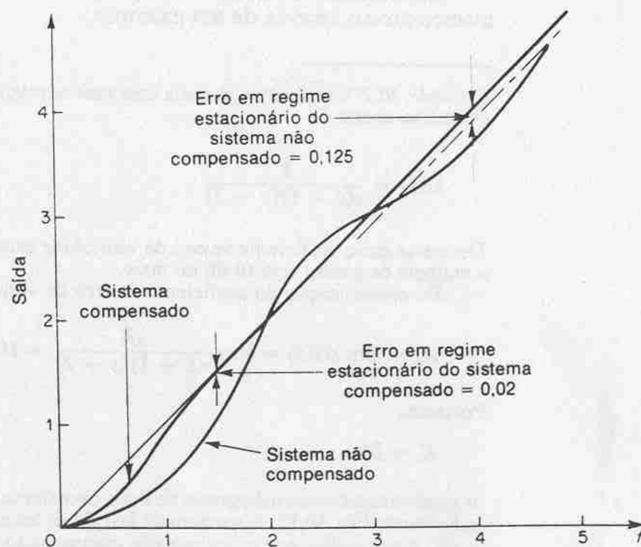


Fig. 10.33 Curvas de resposta transitória para o sistema não compensado e para o sistema compensado. (a) Curvas de resposta ao degrau unitário; (b) curvas de resposta à rampa unitária.

e o sistema compensado possuirá a função de transferência de malha-aberta

$$G_c(s)G(s) = \frac{(s+0,5)(s+0,1)25}{(s+5)(s+0,01)s(s+0,5)}$$

$$= \frac{25(s+0,1)}{s(s+5)(s+0,01)}$$

Devido ao cancelamento dos termos $(s + 0.5)$, o sistema compensado é um sistema de terceira-ordem. (Matematicamente este cancelamento é exato, porém, praticamente, este cancelamento não será exato em virtude de algumas aproximações estarem normalmente envolvidas na dedução do modelo matemático do sistema, e, como resultado, as constantes de tempo não são precisas.) O gráfico do lugar das raízes do sistema compensado é indicado na Fig. 10.32.

Como a contribuição angular da parte de atraso de fase da rede avanço-atraso é muito pequena, não há modificação apreciável na localização dos pólos de malha-fechada dominantes em relação à localização desejada, $s = -2.5 \pm j4.33$. Portanto, o sistema compensado satisfaz todas as especificações de desempenho exigidas. O terceiro pólo de malha-fechada do sistema compensado está localizado em $s = -0,102$. Desde que este pólo de malha-fechada está muito próximo ao zero em $s = -0.01$, o efeito deste pólo na resposta é relativamente pequeno. (Note que, em geral, se um pólo e um zero estão próximos um do outro sobre o eixo real negativo e próximos da origem, então esta combinação de pólo e zero resultará em uma duração longa e de pequena amplitude na resposta transitória.)

As curvas de resposta ao degrau unitário e as curvas de resposta à rampa unitária, antes e após a compensação, são indicadas na Fig. 10.33.

Exemplo 10.6 Considere o sistema de controle do Exemplo 10.5. Suponha que o coeficiente de erro de velocidade estático deve ser 80 s^{-1}. As outras especificações permanecem idênticas às fornecidas no Exemplo 10.5. Isto é, a relação de amortecimento e a freqüência natural não amortecida dos pólos de malha-fechada dominantes são dadas como 0,5 e 5, respectivamente. Projete um compensador avanço-atraso apropriado.

O requisito em relação ao coeficiente de erro de velocidade estático é que $K_v = 80\ s^{-1}$ Este dado resulta em

$$K_c = 10$$

A constante de tempo T_1 e o valor de β são determinados a partir de

$$\left|\frac{s+\dfrac{1}{T_1}}{s+\dfrac{\beta}{T_1}}\right|\left|\frac{40}{s(s+0,5)}\right|_{s=-2,5+j4,33} = \left|\frac{s+\dfrac{1}{T_1}}{s+\dfrac{\beta}{T_1}}\right|\frac{5}{3} = 1$$

$$\left/\frac{s+\dfrac{1}{T_1}}{s+\dfrac{\beta}{T_1}}\right._{s=-2,5+j4,33} = 55°$$

Referindo-se à Fig. 10.34, podemos facilmente localizar os pontos A e B de modo que

$$\angle APB = 55°, \qquad \frac{\overline{PA}}{\overline{PB}} = \frac{3}{5}$$

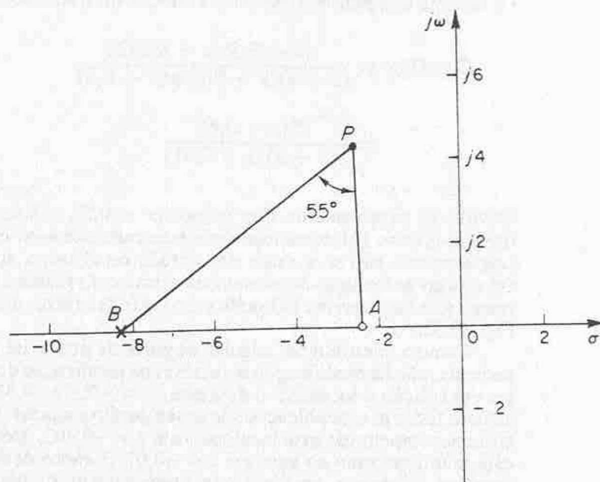


Fig. 10.34 Determinação da localização do pólo e do zero desejados.

O resultado é

$$\overline{AO} = 2{,}4, \qquad \overline{BO} = 8{,}3$$

ou

$$T_1 = \frac{1}{2{,}4} = 0{,}416, \qquad \beta = 8{,}3\,T_1 = 3{,}45$$

A parte do avanço de fase da rede avanço-atraso resulta em

$$\frac{s + 2{,}4}{s + 8{,}3}$$

Para a parte de atraso de fase, podemos escolher

$$T_2 = 10$$

Portanto, o compensador avanço-atraso se torna

$$G_c(s) = \left(\frac{s + 2{,}4}{s + 8{,}3}\right)\left(\frac{s + 0{,}1}{s + 0{,}029}\right)(10)$$

O sistema compensado possuirá a função de transferência de malha-aberta

$$G_c(s)G(s) = \frac{(s + 2{,}4)(s + 0{,}1)40}{(s + 8{,}3)(s + 0{,}029)s(s + 0{,}5)}$$

Neste caso não ocorre cancelamento, e o sistema compensado é de quarta-ordem. Devido à contribuição angular da parte do atraso de fase da rede avanço-atraso ser muito pequena, os

pólos de malha-fechada dominantes estão localizados muito próximos da localização desejada.

Os dois pólos de malha-fechada restantes são obtidos como

$$s = -0{,}09, \qquad s = -3{,}74$$

Desde que o pólo de malha-fechada em $s = -0{,}09$ e o zero em $s = -0{,}1$ praticamente se cancelam, o efeito deste pólo de malha-fechada é muito pequeno. O pólo de malha-fechada restante ($s = -3{,}74$) está relativamente próximo do zero em $s = -2{,}4$, e o efeito deste pólo de malha-fechada na resposta transitória novamente será relativamente pequeno. Os pólos em $s = -2{,}5 \pm j4{,}33$ resultam como pólos de malha-fechada dominantes.

Compensação em avanço-atraso baseada na abordagem da resposta em freqüência. O projeto de um compensador avanço-atraso pela abordagem da resposta em freqüência é baseado na combinação das técnicas de projeto discutidas para a combinação em avanço e para a combinação em atraso. O valor de α para a rede em avanço deve ser igual ao inverso do valor de β para a rede em atraso. Se α é escolhido como $1/\beta$, então podemos simplesmente combinar os compensadores em avanço e em atraso projetados separadamente a fim de obtermos o compensador avanço-atraso único.

A parte do avanço de fase da rede avanço-atraso altera a curva de resposta em freqüência adicionando o ângulo de fase em avanço e aumentando a margem de fase na freqüência de cruzamento do ganho.

A parte do atraso de fase da rede em avanço-atraso acarreta atenuação, próximo e acima da freqüência de cruzamento do ganho, e permite ainda o aumento de ganho na faixa de baixas freqüências a fim de melhorar o desempenho em regime estacionário.

Ilustraremos os detalhes do procedimento projetando um compensador em avanço-atraso através de um exemplo.

Exemplo 10.7 Considere o sistema com realimentação unitária cuja função de transferência em malha-aberta é

$$G(s) = \frac{K}{s(s + 1)(s + 2)}$$

Deseja-se que o coeficiente de erro de velocidade estático seja 10 s^{-1}, a margem de fase 50°, a margem de ganho seja 10 db ou mais.

Da especificação do coeficiente de erro de velocidade estático, obtemos

$$K_v = \lim_{s\to 0} sG(s) = \lim_{s\to 0} \frac{sK}{s(s + 1)(s + 2)} = 10$$

Portanto,

$$K = 20$$

A seguir desenhemos o diagrama de Bode do sistema não compensado com $K = 20$, conforme indicado na Fig. 10.35. A margem de fase do sistema não compensado é determinada, sendo $-32°$, o que indica que o sistema não compensado é instável.

O próximo passo no projeto de um compensador avanço-atraso é escolher uma nova freqüência de cruzamento do ganho. Da curva de ângulo de fase para $G(j\omega)$, note que $\underline{/G(j\omega)} = -180°$ em $\omega = 1{,}5$ rad/s. É conveniente escolher a nova freqüência de cruzamento de ganho em 1,5 rad/s, de modo que o ângulo de avanço de fase exigido em $\omega = 1{,}5$ rad/s é aproximadamente 50°, o que é bem possível pelo uso de uma rede em avanço-atraso única.

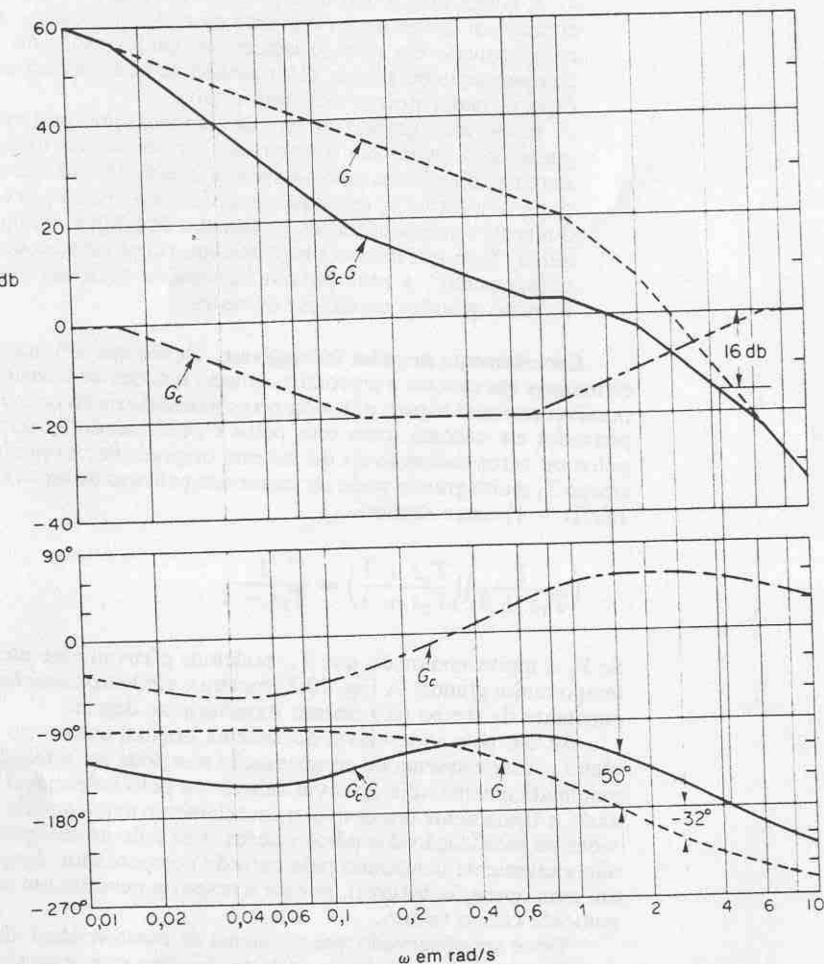


Fig. 10.35 Diagrama de Bode para o sistema não compensado, para o compensador e para o sistema compensado. (G: sistema não compensado, G_c: compensador, G_cG: sistema compensado.)

Uma vez escolhida a freqüência de cruzamento do ganho em 1,5 rad/s, podemos determinar a freqüência de canto da parte de atraso de fase da rede avanço-atraso. Vamos escolher a freqüência de canto $\omega = 1/T_2$ (que corresponde ao zero da parte em atraso de fase da rede) como sendo uma década abaixo da nova freqüência de cruzamento do ganho, ou em $\omega = 0{,}15$ rad/s. Vamos escolher ainda

$$\beta = 10$$

Então, a freqüência de canto $\omega = 1/\beta T_2$ (que corresponde ao pólo da parte em atraso de fase da rede) resulta $\omega = 0{,}015$ rad/s. A função de transferência da parte em atraso de fase da rede

avanço-atraso, então, resulta em

$$\frac{s + 0{,}15}{s + 0{,}015} = 10\left(\frac{6{,}67s + 1}{66{,}7s + 1}\right)$$

A parte de avanço de fase pode ser determinada como segue: Desde que a nova freqüência de cruzamento do ganho está em $\omega = 1{,}5$ rad/s, a partir da Fig. 10.35, $G(j1{,}5)$ é determinado em 13 db. Portanto, se a rede avanço-atraso contribuir com -13 db em $\omega = 1{,}5$ rad/s, então a nova freqüência de cruzamento do ganho é a desejada. Desta especificação, é possível desenhar uma reta com inclinação de 20 db/década passando pelo ponto (-13 db; 1,5 rad/s). As intersecções desta reta e as retas de 0-db e -20 db determinam as freqüências de canto. Portanto, as freqüências de canto para a parte em avanço são $\omega = 0{,}7$ rad/s e $\omega = 7$ rad/s. Conseqüentemente, a função de transferência da parte em avanço da rede avanço-atraso resulta em

$$\frac{s + 0{,}7}{s + 7} = \frac{1}{10}\left(\frac{1{,}43s + 1}{0{,}143s + 1}\right)$$

Combinando as funções de transferência das partes em atraso e em avanço da rede, obtemos a função de transferência do compensador avanço-atraso:

$$\left(\frac{s + 0{,}7}{s + 7}\right)\left(\frac{s + 0{,}15}{s + 0{,}015}\right) = \left(\frac{1{,}43s + 1}{0{,}143s + 1}\right)\left(\frac{6{,}67s + 1}{66{,}7s + 1}\right)$$

As curvas de módulo e ângulo de fase do compensador avanço-atraso que acabamos de projetar são indicadas na Fig. 10.35. A função de transferência em malha-aberta do sistema compensado é

$$G_c(s)G(s) = \frac{(s + 0{,}7)(s + 0{,}15)20}{(s + 7)(s + 0{,}015)s(s + 1)(s + 2)}$$
$$= \frac{10(1{,}43s + 1)(6{,}67s + 1)}{s(0{,}143s + 1)(66{,}7s + 1)(s + 1)(0{,}5s + 1)} \qquad (10.4)$$

As curvas de módulo e ângulo de fase do sistema da Eq. (10.4) também são mostradas na Fig. 10.35. A margem de fase do sistema compensado é 50°, a margem de ganho 16 db, e o coeficiente de erro de velocidade estático é 10 s^{-1}. Todos os requisitos foram, portanto, satisfeitos e o projeto completado.

A Fig. 10.36 mostra os gráficos polares do sistema não compensado e do sistema compensado. O lugar geométrico de $G_c(j\omega)G(j\omega)$ é tangente ao círculo $M = 1{,}2$ em aproximadamente $\omega = 2$ rad/s. Obviamente, isto indica que o sistema compensado possui uma estabilidade relativa satisfatória. A largura de faixa do sistema compensado é ligeiramente maior do que 2 rad/s.

10.6 SUMÁRIO DOS MÉTODOS DE COMPENSAÇÃO DE SISTEMAS DE CONTROLE

Nas Seções 10.3 até 10.5 apresentamos os procedimentos para projeto de compensadores em avanço, em atraso e em avanço-atraso através do uso de exemplos simples. Um projeto satisfatório de um compensador para um dado sistema exigirá a aplicação criativa destes princípios de projeto básicos. (Lembre-se que há uma riqueza de informações na literatura sobre projetos de sistemas de controle complexos, e o leitor deve referir-se a ela para quaisquer problemas específicos que encontrar.)

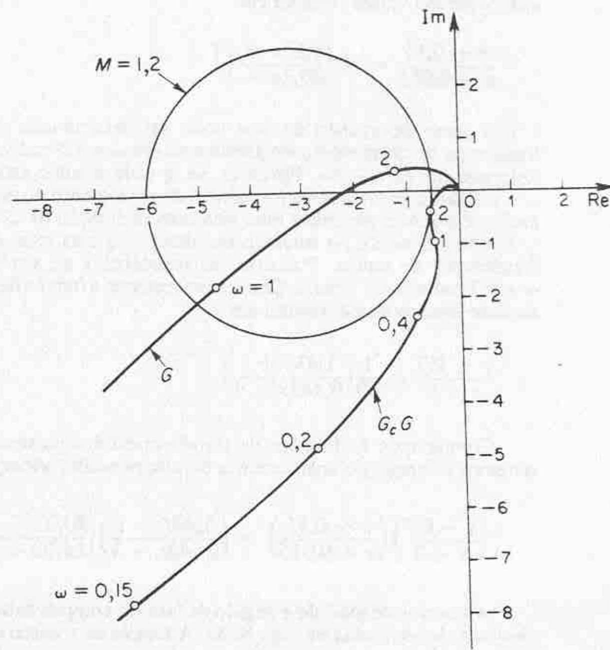


Fig. 10.36 Diagramas polares do sistema não compensado e do sistema compensado. (G: sistema não compensado, G_cG: sistema compensado.)

Comparação das compensações em avanço, em atraso, e em avanço-atraso

1. A compensação em avanço atinge o resultado desejado através dos méritos de sua contribuição em avanço de fase; enquanto que a compensação em atraso atinge o resultado através dos méritos da sua propriedade de atenuação em altas freqüências.
2. No domínio s a compensação em avanço nos possibilita modificar o lugar das raízes e, portanto, propicia os pólos de malha-fechada desejados. No domínio de freqüência, a compensação em avanço aumenta a margem de fase e a largura de faixa. Uma largura de faixa aumentada significa uma redução no tempo de acomodação. A largura de faixa de um sistema com compensação em avanço é sempre maior do que um outro com compensação em atraso. Portanto, se for desejada uma maior largura de faixa, ou resposta mais rápida, deve ser empregada a compensação em avanço. Se, entretanto, estiverem presentes sinais de ruído, então pode não ser desejável uma grande largura de faixa desde que torna o sistema mais susceptível a sinais de ruído, devido ao aumento do ganho em altas freqüências. Neste caso, deve ser usada a compensação em atraso.
3. A compensação em atraso melhora a precisão em regime estacionário; entretanto, reduz a largura de faixa. Se a redução da largura de faixa for excessiva, o sistema compensado apresentará uma resposta lenta. Se forem desejadas tanto resposta rápida como boa precisão estática, deve ser empregado um compensador avanço-atraso.

4. A compensação em avanço exige um aumento adicional no ganho para compensar a atenuação inerente da rede em avanço. Isto significa que a compensação em avanço requer um ganho maior do que o exigido pela compensação em atraso. (Um ganho maior, em muitos casos, implica maior espaço, maior peso e mais alto custo.)
5. Embora um grande número de compensações práticas possam ser conseguidas com redes em avanço, em atraso, ou em avanço-atraso, para um sistema complicado uma compensação simples, através do uso destas redes, pode não fornecer resultados satisfatórios. Então, devem ser empregados diferentes compensadores possuindo diferentes configurações de pólos e zeros. Note que uma vez especificada a configuração de pólos e zeros de um compensador, a rede passiva necessária pode ser realizada pelo uso de técnicas padrões de síntese de redes.

Cancelamento de pólos indesejáveis. Desde que a função de transferência de elementos em cascata é o produto de suas funções de transferência individuais, é possível cancelar alguns pólos ou zeros indesejáveis colocando um elemento compensador em cascata, com seus pólos e zeros sendo ajustados para cancelar os pólos ou zeros indesejáveis do sistema original. Por exemplo, uma constante de tempo T_1 muito grande pode ser cancelada pelo uso de uma rede em avanço $(T_1s + 1)/(T_2s + 1)$ como segue:

$$\left(\frac{1}{T_1s+1}\right)\left(\frac{T_1s+1}{T_2s+1}\right) = \frac{1}{T_2s+1}$$

Se T_2 é muito menor do que T_1, podemos efetivamente eliminar a constante de tempo muito grande. A Fig. 10.37 mostra o efeito do cancelamento de uma grande constante de tempo na resposta transitória ao degrau.

Se um pólo indesejável do sistema original estiver no semiplano direito do plano s, este esquema de compensação não pode ser utilizado desde que, embora matematicamente seja possível cancelar o pólo indesejável com um zero adicionado, é fisicamente impossível o cancelamento exato devido às imprecisões envolvidas na localização dos pólos e zeros. Um pólo no semiplano direito do plano s, não exatamente cancelado pelo zero do compensador, eventualmente redundará em uma operação instável, porque a resposta possuirá um termo exponencial que aumenta com o tempo.

Deve ser observado que o sistema de controle ideal não é aquele que possui uma função de transferência unitária. Fisicamente, este sistema de controle não pode ser construído, já que não se pode transferir energia instantaneamente da entrada para a saída. Além disso, desde que ruído está sempre presente de uma forma ou de outra, um sistema com uma função de transferência unitária não é desejável. Um sistema de controle desejado, em muitos casos práticos, pode

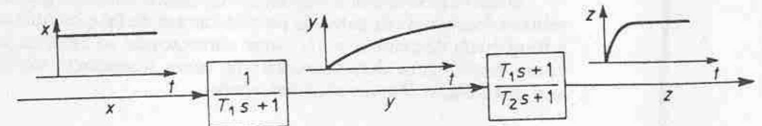


Fig. 10.37 Curvas de resposta ao degrau mostrando o efeito do cancelamento de uma constante de tempo grande.

possuir um conjunto de pólos conjugados dominantes de malha-fechada com uma relação de amortecimento e uma freqüência natural não amortecida razoáveis. A determinação da parte significativa da configuração de pólos e zeros de malha-fechada, tal como a localização dos pólos de malha-fechada dominantes, é baseada nas especificações que fornecem o desempenho desejado do sistema.

Cancelamento de pólos complexos conjugados indesejáveis. Se a função de transferência de um processo contém um ou mais pares de pólos complexos conjugados, então um compensador em avanço, atraso ou avanço-atraso pode não fornecer resultados satisfatórios. Neste caso, uma rede que possui dois zeros e dois pólos pode ser útil. Se os zeros forem escolhidos de modo a cancelar os pólos complexos conjugados indesejáveis do processo, podemos, essencialmente, substituir os pólos indesejáveis por aqueles aceitáveis. Isto é, se os pólos complexos conjugados indesejáveis estiverem no semiplano esquerdo do plano s e possuírem a forma

$$\frac{1}{s^2 + 2\zeta_1\omega_1 s + \omega_1^2}$$

então a inserção de uma rede compensadora possuindo a função de transferência

$$\frac{s^2 + 2\zeta_1\omega_1 s + \omega_1^2}{s^2 + 2\zeta_2\omega_2 s + \omega_2^2}$$

resultará em uma mudança efetiva dos pólos complexos conjugados para aqueles aceitáveis. Note que mesmo que o cancelamento não seja exato, o sistema compensado possuirá melhores características de resposta. (Conforme estabelecido anteriormente, esta abordagem não pode ser utilizada se os pólos complexos conjugados indesejáveis estiverem no semiplano direito do plano s.)

Redes familiares que consistem apenas em componentes RC cujas funções de transferência possuem dois zeros e dois pólos são as redes em ponte tipo T. Exemplos de redes em ponte tipo T, e suas funções de transferências, são fornecidos na Fig. 10.38.

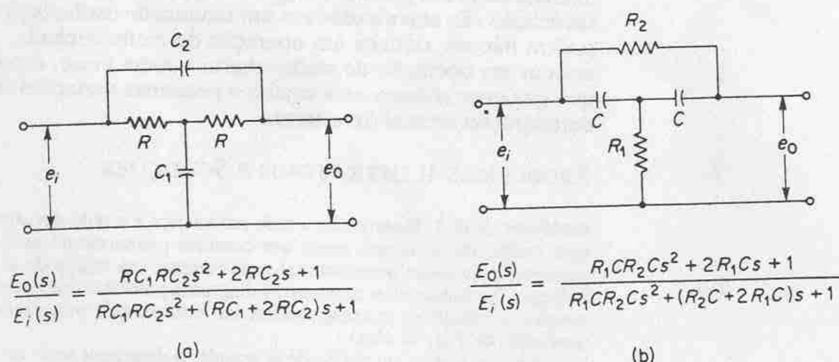
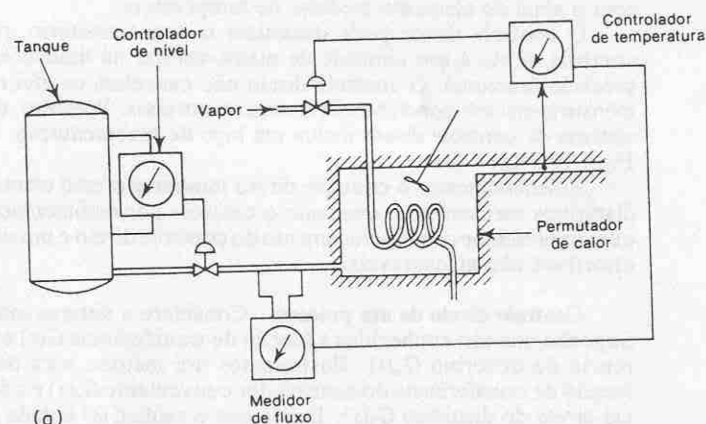


Fig. 10.38 Redes em ponte tipo T.

Compensação por realimentação. A realimentação tacométrica é muito comumente utilizada em servomecanismos de posição. Na Seção 6.4 discutimos um problema de projeto simples utilizando a realimentação tacométrica. Portanto, não o repetiremos aqui.

Um tacômetro é um dos dispositivos denominados dispositivos de realimentação por taxa. Um outro dispositivo de realimentação por taxa muito comum é a taxa de rotação. Dispositivos de taxa de rotação são comumente utilizados em sistemas de aeronaves com piloto automático.

Eliminação de efeitos indesejáveis de distúrbios por controle no ramo direto. Se os distúrbios forem mensuráveis, o controle no ramo direto é um método útil de cancelamento de seus efeitos na saída do sistema. Por controle no ramo direto, designamos o controle de efeitos indesejáveis de distúrbios mensuráveis através de compensação aproximada dos mesmos antes que eles se materializem. Isto é vantajoso porque em um sistema de controle com realimentação usual, a ação corretiva somente se inicia após a saída haver sido afetada.



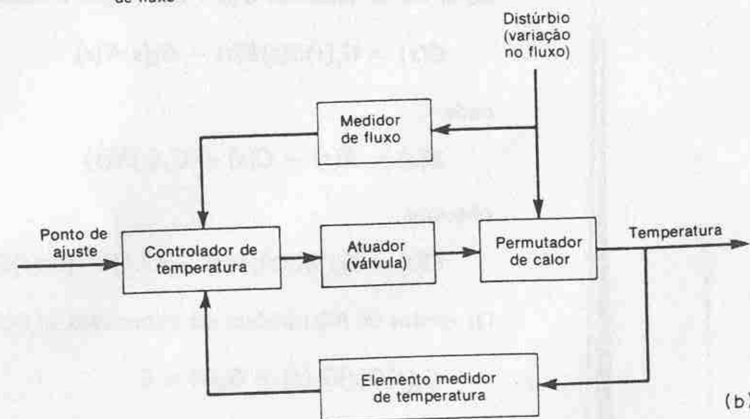


Fig. 10.39 (a) Sistema de controle de temperatura; (b) diagrama de blocos.

Como um exemplo, considere o sistema de controle de temperatura indicado na Fig. 10.39(a). Neste sistema, deseja-se manter a temperatura de saída em um valor constante. O distúrbio neste sistema é uma variação na taxa de fluxo de entrada, que depende do nível no tanque. O efeito de uma variação nesta taxa não pode ser sentido imediatamente na saída devido ao atraso de tempo envolvido no sistema.

O controlador de temperatura, que controla a entrada de calor para o permutador térmico, não atuará até perceber o erro. Se o sistema possui grandes tempos de atraso, transcorrerá um certo intervalo de tempo antes de haver qualquer ação corretiva. De fato, quando o erro for percebido após um certo atraso e a ação corretiva começar, pode ser muito tarde para manter a temperatura de saída dentro dos limites desejados.

Se for utilizado o controle direto neste sistema, então, logo que ocorrer uma variação no fluxo de entrada, haverá simultaneamente uma medida corretiva, ajustando a entrada de calor através do permutador de calor. Isto pode ser realizado alimentando-se o controlador de temperatura tanto com o sinal do fluxímetro como com o sinal do elemento medidor de temperatura.

O controle direto pode minimizar o erro transitório, porém, desde que o controle direto é um controle de malha-aberta, há limitações em relação à sua precisão funcional. O controle direto não cancelará os efeitos de distúrbios não mensuráveis sob condições de operação normais. Portanto, é necessário que um sistema de controle direto inclua um laço de realimentação, como mostrado nas Figs. 10.39(a) e (b).

Essencialmente, o controle direto minimiza o erro transitório causado pelos distúrbios mensuráveis, enquanto o controle por realimentação compensa quaisquer imperfeições no funcionamento do controle direto e possibilita correções para distúrbios não mensuráveis.

Controle direto de um processo. Considere o sistema indicado na Fig. 10.40. Suponha que são conhecidas a função de transferência $G(s)$ e a função de transferência do distúrbio $G_n(s)$. Ilustraremos um método para determinação de uma função de transferência do controlador conveniente $G_c(s)$ e a função de transferência direta do distúrbio $G_1(s)$. Desde que a saída $C(s)$ é dada por

$$C(s) = G_c(s)G(s)E(s) + G_n(s)N(s)$$

onde

$$E(s) = R(s) - C(s) + G_1(s)N(s)$$

obtemos

$$C(s) = G_c(s)G(s)[R(s) - C(s)] + [G_c(s)G(s)G_1(s) + G_n(s)]N(s)$$

Os efeitos de $N(s)$ podem ser eliminados se $G_c(s)$ é escolhida de modo que

$$G_c(s)G(s)G_1(s) + G_n(s) = 0$$

ou

$$G_1(s) = -\frac{G_n(s)}{G_c(s)G(s)} \tag{10.5}$$

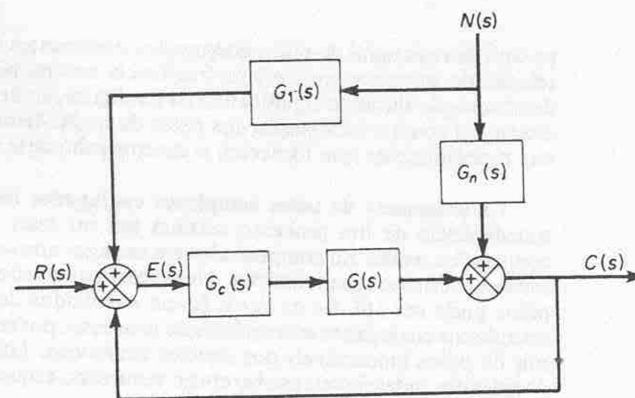


Fig. 10.40 Sistema de controle.

Projetando apropriadamente a função de transferência do controlador $G_c(s)$ (conforme discutido nas Seções 10.3 até 10.5), o sistema de controle de malha-fechada pode apresentar o desempenho desejado. Uma vez determinada $G_c(s)$, então a função de transferência direta do distúrbio $G_1(s)$ pode ser obtida da Eq. (10.5).

Comentários conclusivos. Nos exemplos de projeto apresentados neste capítulo, consideramos principalmente e apenas as funções de transferência de compensadores. Nos problemas de projeto reais, devemos determinar os componentes. Portanto, devemos satisfazer limitações de projeto adicionais tais como custo, dimensão, peso e confiabilidade.

O sistema projetado pode satisfazer as especificações sob condições de operação normais, porém pode desviar-se consideravelmente das especificações quando ocorrerem variações ambientais consideráveis. Desde que as variações ambientais afetam o ganho e as constantes de tempo do sistema, é necessário propiciar meios automáticos ou manuais para ajustar o ganho a fim de compensar estas variações ambientais, os efeitos não lineares que não foram levados em consideração no projeto, e também para compensar tolerâncias de fabricação de unidade para unidade na produção dos componentes do sistema. (Os efeitos das tolerâncias de fabricação são amenizados em um sistema de malha-fechada; portanto, os efeitos podem não ser críticos em operação de malha-fechada, porém podem tornar-se críticos em operação de malha-aberta.) Além disso, o projetista deve lembrar-se que qualquer sistema está sujeito a pequenas variações devidas principalmente à deterioração normal do sistema.

Problemas Ilustrativos e Soluções

Problema A.10.1 Mostre que a rede em avanço e a rede em atraso inseridas em cascata em uma malha-aberta atuam como um controle proporcional-mais-derivativo (na região de ω pequeno) e controle proporcional-mais-integral (na região de ω grande), respectivamente.

Solução. Na região de ω pequeno, o diagrama polar da rede em avanço é aproximadamente o mesmo do diagrama correspondente ao controlador proporcional-mais-derivativo. Isto é mostrado na Fig. 10.41(a).

Analogamente, na região de ω grande, o diagrama polar da rede em atraso se aproxima daquele do controlador proporcional-mais-integral, conforme indicado na Fig. 10.41(b).

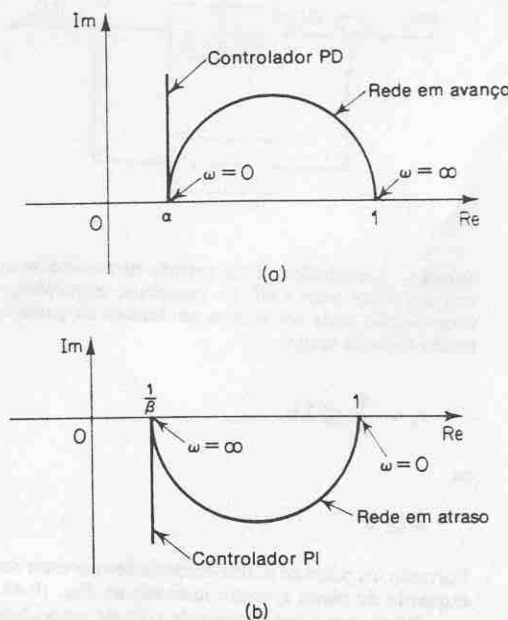


Fig. 10.41 (a) Diagramas polares de uma rede em avanço e de um controlador proporcional-mais-derivativo; (b) diagramas polares de uma rede em atraso e um controlador proporcional-mais-integral.

Problema A.10.2 Se a função de transferência de malha-aberta $G(s)$ possui pólos complexos conjugados ligeiramente amortecidos, então, mais do que um lugar geométrico M pode ser tangente ao lugar geométrico de $G(j\omega)$.

Considere o sistema com realimentação unitária cuja função de transferência de malha-aberta é

$$G(s) = \frac{9}{s(s + 0{,}5)(s^2 + 0{,}6s + 10)} \tag{10.6}$$

Desenhe os diagramas de Bode para esta função de transferência de malha-aberta. Desenhe também o gráfico log-módulo *versus* fase e mostre que dois lugares geométricos M são tangentes ao lugar geométrico de $G(j\omega)$. Finalmente, desenhe o gráfico dos diagramas de Bode para a função de transferência de malha-fechada.

Solução. A Fig. 10.42 mostra os diagramas de Bode de $G(j\omega)$. A Fig. 10.43 mostra o gráfico do log-módulo *versus* fase de $G(j\omega)$. Pode ser visto que o lugar geométrico de $G(j\omega)$ é tangente ao lugar geométrico de $M = 8$ db em $\omega = 0{,}97$ rad/s e tangente ao lugar geométrico de $M = -4$ db em $\omega = 2{,}8$ rad/s.

A Fig. 10.44 mostra o diagrama de Bode da função de transferência de malha-fechada. As curvas de resposta em freqüência de malha-fechada indicam dois picos de ressonância. Note que este fato ocorre quando a função de transferência de malha-fechada envolve o produto de dois termos de segunda-ordem ligeiramente amortecidos e as duas freqüências de ressonância

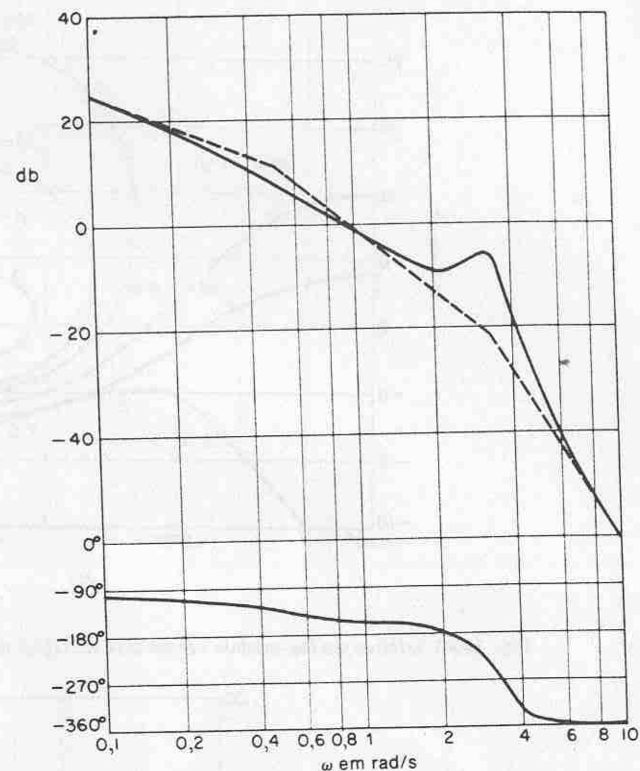


Fig. 10.42 Diagrama de Bode de $G(j\omega)$ dado pela Eq. (10.6).

correspondentes estão suficientemente separadas uma da outra. De fato, a função de transferência de malha-fechada deste sistema pode ser escrita

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)}$$

$$= \frac{9}{(s^2 + 0{,}487s + 1)(s^2 + 0{,}613s + 9)}$$

Obviamente, a função de transferência de malha-fechada é um produto de dois termos de segunda-ordem ligeiramente amortecidos (as relações de amortecimento são 0,24 e 0,1) e as duas freqüências de ressonância estão suficientemente separadas.

Problema A.10.3 Considere o sistema mostrado na Fig. 10.45. Determine os valores do ganho K do amplificador e do ganho K_h da realimentação de velocidade de modo a serem satisfeitas as seguintes especificações:

1. relação de amortecimento dos pólos de malha-fechada é 0,5.
2. tempo de acomodação ≤ 2 s.
3. coeficiente de erro de velocidade $K_v \geq 50\ \text{s}^{-1}$.
4. $0 < K_h < 1$.

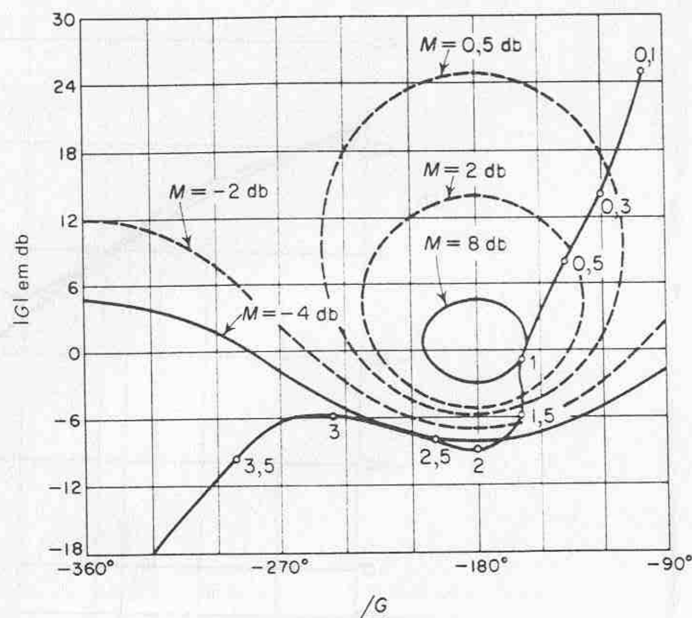


Fig. 10.43 Gráfico do log-módulo *versus* fase de $G(j\omega)$ dada pela Eq. (10.6).

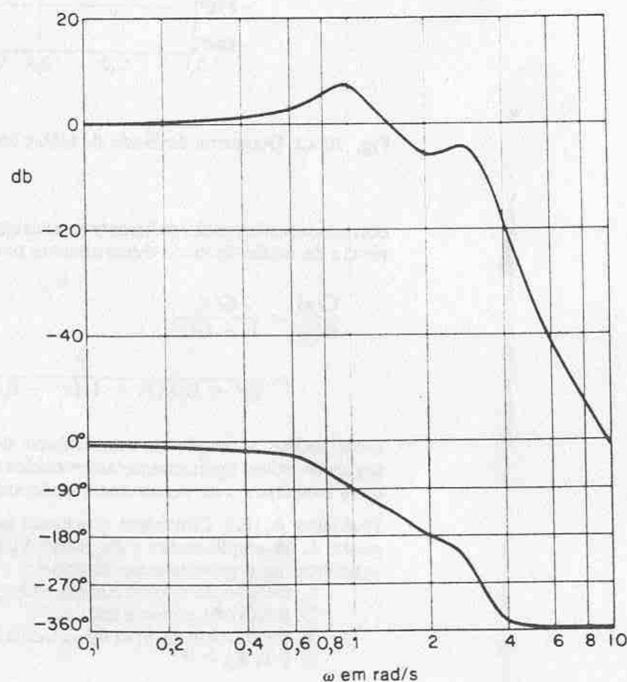


Fig. 10.44 Diagrama de Bode de $G(j\omega)/[1 + G(j\omega)]$ onde $G(j\omega)$ é dada pela Eq. (10.6).

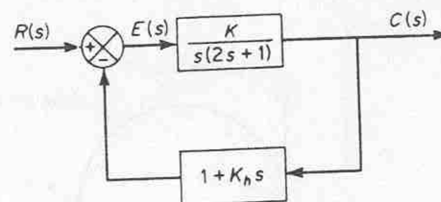


Fig. 10.45 Sistema de controle.

Solução. A especificação da relação de amortecimento exige que os pólos de malha-fechada estejam sobre retas a 60º no semiplano esquerdo do plano s. A especificação do tempo de acomodação pode ser escrita em termos da parte real dos pólos complexos conjugados de malha-fechada como

$$t_s = \frac{4}{\sigma} \leq 2\text{ s}$$

ou

$$\sigma \geq 2$$

Portanto, os pólos de malha-fechada devem estar sobre as retas cheias AB e CD no semiplano esquerdo do plano s, como indicado na Fig. 10.46.

Desde que o coeficiente de erro de velocidade K_v é definido por

$$K_v = \lim_{s \to 0} sG(s)H(s)$$

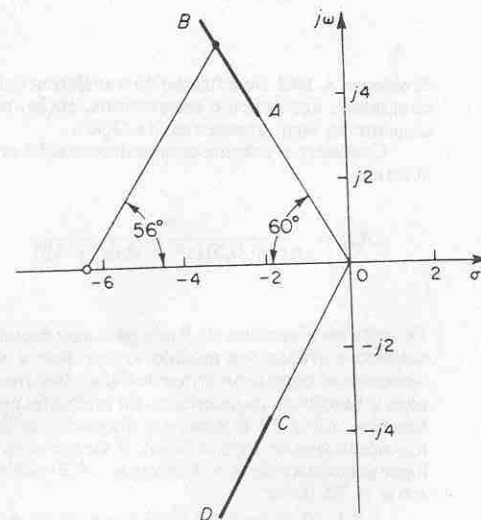


Fig. 10.46 Localizações possíveis para os pólos de malha-fechada no plano s para o sistema do Problema A.10.3.

obtemos

$$K_v = \lim_{s \to 0} \frac{sK(1 + K_h s)}{s(2s+1)} = K$$

Da especificação fornecida relativa ao coeficiente de erro de velocidade, obtemos

$$K \geq 50$$

Para este sistema, os pólos de malha-aberta estão em $s = 0$ e $s = -1/2$. O zero de malha-aberta está em $s = -1/K_h$, onde K_h é uma constante indeterminada. Inicialmente, consideremos que os pólos de malha-fechada estejam em $s = -2 \pm j3{,}4$ (pontos A e C na Fig. 10.46). A soma dos ângulos, na localização do pólo de malha-fechada escolhido, com os pólos de malha-aberta é $120° + 115° = 235°$. Portanto, necessitamos de uma contribuição de 55° do zero, de modo que a soma total seja $-180°$. Para satisfazer a condição do ângulo, escolhemos o zero em $s = -4{,}4$. Então K_h é

$$K_h = \frac{1}{4{,}4} = 0{,}227$$

A condição de módulo estabelece que

$$\left| \frac{K(1 + 0{,}227s)}{s(2s+1)} \right|_{s=-2+j3{,}4} = 1$$

Portanto,

$$K = 31$$

Desde que $K < 50$, a escolha dos pólos de malha-fechada em $-2 \pm j3{,}4$ não é aceitável.

Como uma segunda tentativa, consideremos os pólos de malha-fechada em $-3 \pm j5{,}1$. A soma das contribuições angulares dos pólos de malha-aberta é 236°. Necessitamos de uma contribuição de 56° para o zero. Isto implica que o zero deve estar em $s = -6{,}4$.

A aplicação da condição do módulo fornece $K = 70$. Este resultado é bastante satisfatório. Desde que $K_h = 0{,}156$, está satisfeita a exigência em relação a K_h. Portanto, estão satisfeitas todas as especificações. Um conjunto de valores aceitáveis para K e K_h, portanto, é o seguinte:

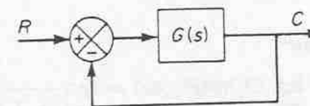
$$K = 70,$$

$$K_h = 0{,}156$$

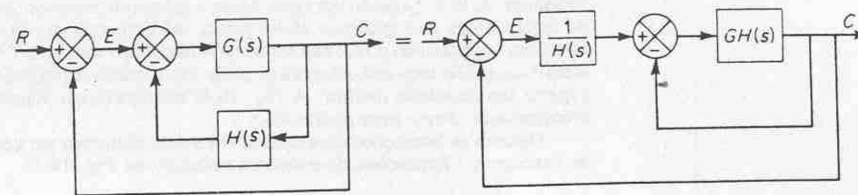
Problema A.10.4 Um sistema de malha-fechada possui a propriedade de que sua função de transferência de malha-fechada é aproximadamente igual ao inverso da função de transferência do ramo de realimentação quando o ganho de malha-aberta é muito maior do que a unidade.

A característica de malha-aberta pode ser modificada adicionando-se uma malha de realimentação interna com uma característica igual ao inverso da característica de malha-aberta desejada. Suponha que um sistema com realimentação unitária possua a seguinte função de transferência de malha-aberta:

$$G(s) = \frac{K}{(T_1 s + 1)(T_2 s + 1)}$$



(a)



(b)

Fig. 10.47 (a) Sistema de controle; (b) adição da malha de realimentação interna para modificar a característica de malha-fechada.

Determine a função de transferência $H(s)$ do elemento na malha de realimentação interna de modo que a malha interna resulte ineficaz tanto em baixas como em altas freqüências.

Solução. A Fig. 10.47(a) mostra o sistema original. A Fig. 10.47(b) mostra a adição da malha de realimentação interna em volta de $G(s)$. Desde que

$$\frac{C(s)}{E(s)} = \frac{G(s)}{1 + G(s)H(s)} = \frac{1}{H(s)} \frac{G(s)H(s)}{1 + G(s)H(s)}$$

se o ganho ao longo da malha interna é muito maior do que a unidade, então $G(s)H(s)/[1 + G(s)H(s)]$ é aproximadamente igual à unidade e a função de transferência $C(s)/E(s)$ é aproximadamente igual a $1/H(s)$.

Por outro lado, se o ganho $G(s)H(s)$ é muito menor do que a unidade, a malha interna resulta ineficaz e $C(s)/E(s)$ é aproximadamente igual a $G(s)$.

Para tornar a malha interna ineficiente tanto na faixa de baixas como na faixa de altas freqüências, necessitamos que

$$G(j\omega)H(j\omega) \ll 1 \quad \text{para } \omega \ll 1 \quad \text{e} \quad \omega \gg 1$$

Desde que neste problema

$$G(j\omega) = \frac{K}{(1 + j\omega T_1)(1 + j\omega T_2)}$$

a exigência pode ser satisfeita se $H(s)$ for escolhida como sendo

$$H(s) = ks$$

porque

$$\lim_{\omega\to 0} G(j\omega)H(j\omega) = \lim_{\omega\to 0} \frac{Kkj\omega}{(1 + j\omega T_1)(1 - j\omega T_2)} = 0$$

$$\lim_{\omega\to \infty} G(j\omega)H(j\omega) = \lim_{\omega\to \infty} \frac{Kkj\omega}{(1 + j\omega T_1)(1 - j\omega T_2)} = 0$$

Portanto, com $H(s) = ks$, o laço interno torna-se ineficiente tanto nas regiões de baixas freqüências como nas regiões de altas freqüências.

Problema A.10.5 Quando um distúrbio age sobre um processo, transcorre um certo intervalo de tempo antes que qualquer efeito possa ser detectado na saída. Se medirmos o próprio distúrbio (embora isto possa não ser possível ou possa ser muito difícil) ao invés da resposta ao distúrbio, então uma ação corretiva pode ser tomada rapidamente, mais cedo, e podemos esperar um resultado melhor. A Fig. 10.48 apresenta um diagrama de blocos mostrando a compensação direta para o distúrbio.

Discuta as limitações do esquema direto do distúrbio em geral. Discuta posteriormente as vantagens e limitações do esquema indicado na Fig. 10.48.

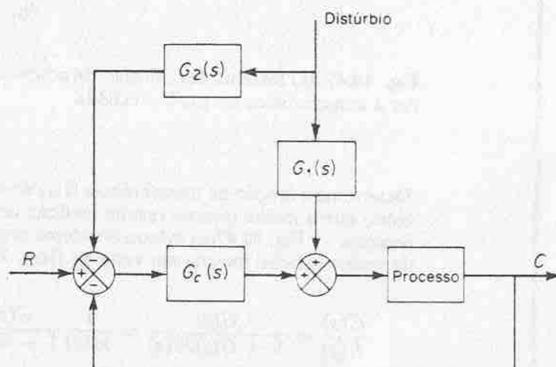


Fig. 10.48 Sistema de controle com compensação direta para o distúrbio.

Solução. Um esquema direto do distúrbio é um esquema de malha-aberta e portanto dependente da constância dos valores dos parâmetros. Qualquer deriva nestes valores resultará em uma compensação imperfeita.

No sistema aqui apresentado, os esquemas de malha-aberta e de malha-fechada estão em operação simultaneamente. Os grandes erros, devidos à fonte de distúrbios principal, podem ser muito reduzidos pela compensação em malha-aberta sem exigir um alto ganho de malha. Erros menores devidos a outras fontes de distúrbio podem ser levados em consideração pelo esquema de controle de malha-fechada. Portanto, erros devidos a todas as origens podem ser reduzidos sem exigir um alto ganho de malha. Esta é uma vantagem do ponto de vista de estabilidade.

Note que este esquema não pode ser usado a menos que o próprio distúrbio principal possa ser medido.

PROBLEMAS

Problema B.10.1 Desenhe os diagramas de Bode da rede em avanço e da rede em atraso indicadas nas Figs. 10.49(a) e (b), respectivamente.

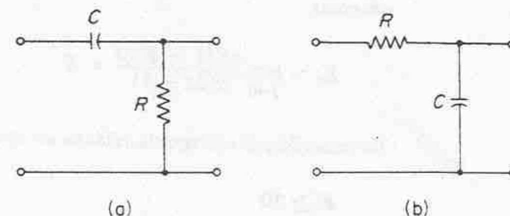


Fig. 10.49 (a) Rede em avanço; (b) rede em atraso.

Problema B.10.2 Considere um sistema de controle com realimentação unitária cuja função de transferência do ramo direto é dada por

$$G(s) = \frac{K}{s(s + 1)(s + 2)(s + 3)}$$

Determine o valor de K de modo que os pólos de malha-fechada dominantes possuam uma relação de amortecimento = 0.5.

Problema B.10.3 Considere um sistema com realimentação unitária cuja função de transferência do ramo direto é dada por

$$G(s) = \frac{1}{s^2}$$

Deseja-se inserir um compensador série de modo que a curva de resposta em freqüência de malha-aberta seja tangente ao círculo M = 3 db em ω = 3 rad/s. O sistema está sujeito a ruídos de alta freqüência e, portanto, deseja-se um corte agudo. Projete um compensador série apropriado.

Problema B.10.4 Determine os valores de K, T_1 e T_2 do sistema mostrado na Fig. 10.50 de modo que os pólos dominantes de malha-fechada tenham $\zeta = 0.5$ e $\omega_n = 3$ rad/s.

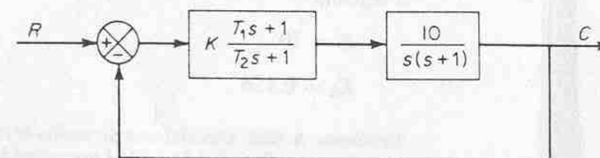


Fig. 10.50 Sistema de controle.

Problema B.10.5 Considere um sistema de controle com realimentação unitária cuja função de transferência do ramo direto é dada por

$$G(s) = \frac{10}{s(s + 2)(s + 8)}$$

Projete um compensador tal que o coeficiente de erro de velocidade estático K_v seja igual a 80 s^{-1} e os pólos dominantes de malha-fechada estejam localizados em $s = -2 \pm j2\sqrt{3}$.

Problema B.10.6 Considere o sistema indicado na Fig. 10.51. Se o distúrbio N puder ser detectado, pode-se passá-lo através de uma função de transferência G_3 e adicioná-lo ao ramo direto entre o amplificador e o processo, conforme indicado na Fig. 10.51.

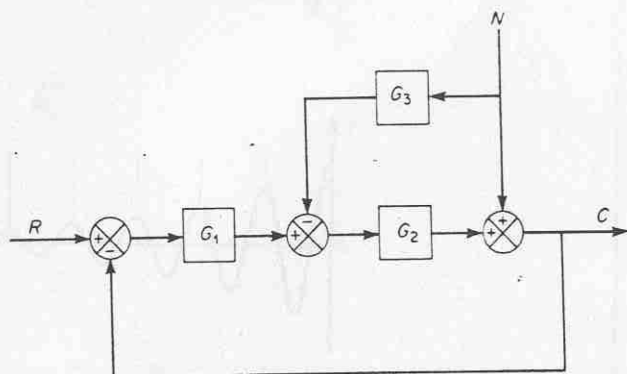


Fig. 10.51 Sistema de controle.

Para reduzir o efeito deste distúrbio N no erro de regime estacionário, determine uma função de transferência G_3 apropriada. O que limitará a presente abordagem na redução dos efeitos deste distúrbio?

Problema B.10.7 A Fig. 10.52 apresenta o diagrama de blocos de um sistema de controle de taxa de atitude. Projete um compensador $G_c(s)$ de modo que os pólos complexos conjugados (dominantes) estejam em $s = -2 \pm j2$.



Fig. 10.52 Sistema de controle de taxa de atitude.

11

Análise de Sistemas de Controle não Lineares por Função Descritiva

11.1 INTRODUÇÃO A SISTEMAS NÃO LINEARES

Já é bem sabido que muitas relações entre quantidades físicas não são muito lineares, embora freqüentemente sejam aproximadas por equações lineares, principalmente pela simplicidade matemática. Esta simplificação pode ser satisfatória desde que as soluções resultantes estejam de acordo com os resultados experimentais. Uma das características mais importantes de sistemas não lineares é a dependência do comportamento da resposta do sistema em relação à amplitude e tipo da entrada. Por exemplo, um sistema não linear pode-se comportar de forma completamente diferente em resposta a entradas em degrau de diferentes amplitudes.

Conforme mencionamos no Cap. 4, os sistemas não lineares diferem dos sistemas lineares em que o princípio da superposição não vale para os não lineares. Os sistemas não lineares apresentam muitos fenómenos que não podem ser vistos em sistemas lineares, e ao investigar tais sistemas devemos estar familiarizados com estes fenómenos.

Nesta seção, iremos apresentar uma breve discussão de vários dos fenómenos.

Dependência freqüência-amplitude. Considere a oscilação natural ou livre do sistema mecânico visto na Fig. 11.1, que consiste em uma massa, um amortecedor viscoso, e uma mola não linear. A equação diferencial que descreve a dinâmica deste sistema pode ser escrita como

$$m\ddot{x} + f\dot{x} + kx + k'x^3 = 0 \qquad (11.1)$$

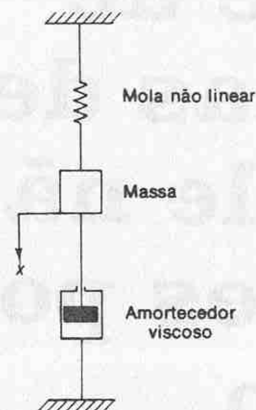


Fig. 11.1 Sistema mecânico.

onde

x = deslocamento da massa
m = massa
f = coeficiente de atrito viscoso do amortecedor
$kx + k'x^3$ = força da mola não linear

Os parâmetros m, f e k são constantes positivas, enquanto que k' pode ser tanto positivo como negativo. Se k' é positivo, a mola é chamada uma mola dura (*hard spring*); se k' é negativo, mola macia. O grau de não linearidade do sistema é caracterizado pela magnitude de k'. Esta equação diferencial não linear, Eq. (11.1), é conhecida como a equação de Duffing e tem sido freqüentemente discutida no campo da mecânica não linear. A solução da Eq. (11.1) representa uma oscilação amortecida se o sistema é sujeito a condições iniciais não nulas. Em uma investigação experimental, observa-se que quando a amplitude decresce, a freqüência da oscilação natural ou decresce ou cresce, dependendo se $k' > 0$ ou $k' < 0$, respectivamente. Quando $k' = 0$, a freqüência permanece inalterada quando a amplitude da oscilação natural decresce. (Isto corresponde a um sistema linear.) Estas características são vistas na Fig. 11.2, que mostra as formas de onda das oscilações naturais. A Fig. 11.3 mostra relação de amplitude-freqüência para os três casos em que k' é maior, igual, ou menor que zero.

Em um estudo experimental de sistemas não lineares, a dependência amplitude-freqüência pode ser facilmente detectada. A dependência amplitude-freqüência é uma das características mais fundamentais de oscilações em sistemas não lineares. Um gráfico da forma visto na Fig. 11.3 revela se uma não linearidade está presente e também indica o grau de não linearidade.

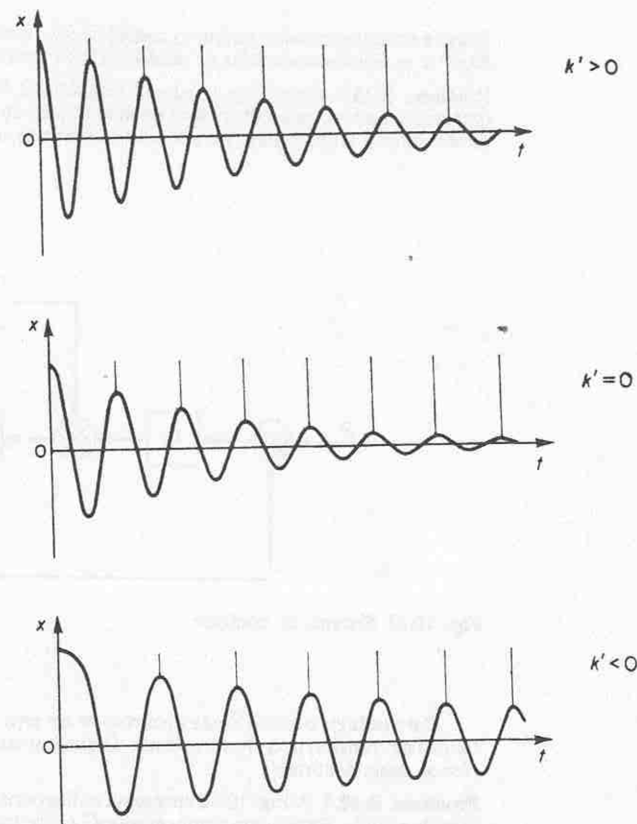


Fig. 11.2 Formas de onda de oscilações naturais do sistema descrito pela Eq. (11.1).

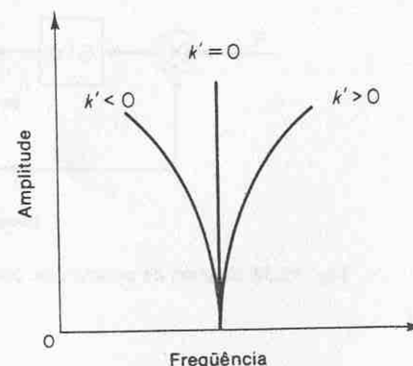


Fig. 11.3 Curvas de amplitude em função da freqüência para oscilações naturais para o sistema descrito pela Eq. (11.1).

Respostas com múltiplos valores e ressonâncias com saltos. Ao se fazer experimentos com oscilações forçadas no sistema visto na Fig. 11.1, cuja equação diferencial é

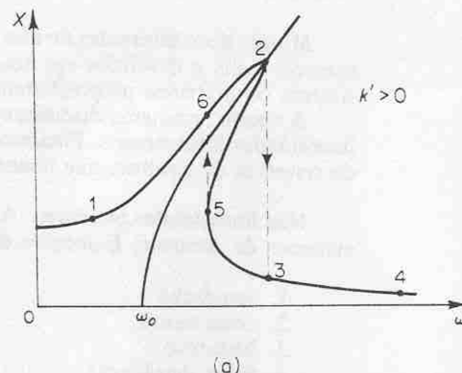
$$m\ddot{x} + f\dot{x} + kx + k'x^3 = P\cos\omega t$$

onde

$P\cos\omega t$ = função de excitação

podemos observar vários fenômenos, tais como respostas com múltiplos valores, ressonâncias com saltos e uma variedade de movimentos periódicos (tais como oscilações sub-harmônicas e oscilações super-harmônicas). Estes fenômenos não ocorrem nas respostas de sistemas lineares.

Ao realizar experimentos em que a amplitude P da função de excitação é mantida constante, enquanto sua freqüência é variada lentamente e a amplitude X da resposta observada, podemos obter uma curva de resposta em freqüência semelhante às vistas nas Figs. 11.4(a) e (b). Suponha que $k' > 0$ e que a freqüência de entrada ω é baixa no começo no ponto 1 da curva da Fig. 11.4(a). À medida que a freqüência ω é aumentada, a amplitude X aumenta até que o ponto 2 é alcançado. Um aumento adicional na freqüência ω causará um salto do ponto 2 ao ponto 3, acompanhados de mudanças na amplitude e fase. Este fenômeno é a chamada ressonância com saltos. Quando a freqüência ω é aumentada ainda mais, a amplitude X segue a curva do ponto 3 em direção ao ponto 4. Ao se fazer o experimento no outro sentido, isto é, começando de uma freqüência alta, observamos que quando ω é diminuído, a amplitude X cresce devagar passando pelo ponto 3, até que o ponto 5 é alcançado. Um decréscimo adicional em ω causará um outro salto do ponto 5 ao ponto 6, acompanhado por mudanças de amplitude e fase. Após este salto, a amplitude X diminui com ω e segue a curva do ponto 6 em direção ao ponto 1. Portanto, as curvas de resposta são na realidade descontínuas, e um ponto representativo na curva de resposta segue caminhos diferentes para freqüências crescentes e decrescentes. As oscilações de resposta correspondentes à curva entre o ponto 2 e o ponto 5 correspondem a oscilações instáveis e não podem ser observadas experimentalmente. Saltos semelhantes acontecem no caso de um sistema com uma mola macia ($k' < 0$), como visto na Fig. 11.4(b). Portanto vemos que para uma dada amplitude P da função de excitação, há uma gama de freqüências em que pode ocorrer uma das duas oscilações estáveis. Deve-se notar que para a existência de ressonância com saltos é necessário que o termo de amortecimento seja pequeno e que a amplitude da função de excitação seja suficientemente grande para levar o sistema para uma região de operação consideravelmente não linear.



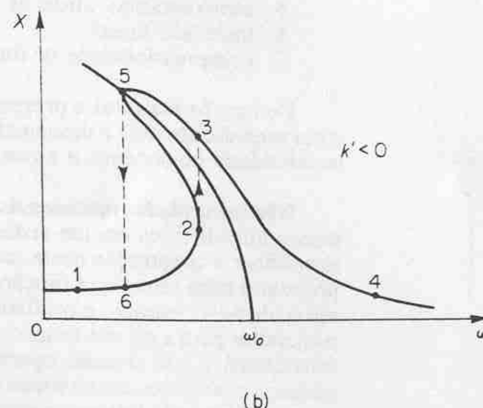


Fig. 11.4 Curvas de resposta em freqüência mostrando ressonâncias com saltos. (a) Sistema mecânico com mola dura; (b) sistema mecânico com mola macia.

Oscilações sub-harmônicas. Chamamos de *oscilações sub-harmônicas* as oscilações não lineares estacionárias cujas freqüências são um inteiro submúltiplo da freqüência de excitação. Um exemplo de uma forma de onda em oscilação sub-harmônica é visto na Fig. 11.5, juntamente com a forma de onda da entrada. A geração de oscilações sub-harmônicas é dependente de parâmetros do sistema e de condições iniciais, bem como da amplitude e freqüência da função de excitação. A frase *dependência de condições iniciais* significa que as oscilações sub-harmônicas não começam por si só. É necessário dar uma espécie de "empurrão", por exemplo, uma variação abrupta da amplitude ou freqüência da função de excitação para iniciar tais oscilações. Uma vez que oscilações sub-harmônicas são excitadas, elas podem ser bastante estáveis para certa faixa de freqüências. Se a freqüência da função de excitação é mudada para um novo valor, ou as oscilações sub-harmônicas desaparecem ou a freqüência da oscilação sub-harmônica também muda para um valor que é ω/n, onde ω é a freqüência de excitação e n é a ordem da oscilação sub-harmônica. (Note que uma oscilação cuja freqüência é a metade daquela da

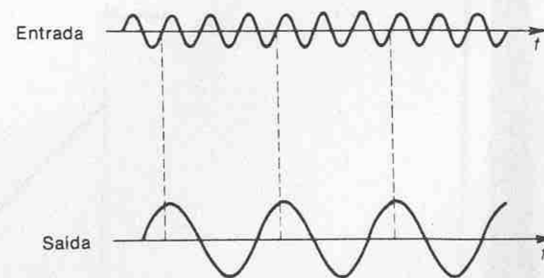


Fig. 11.5 Formas de onda de entrada e saída sob oscilação sub-harmônica.

função de excitação pode acontecer em um sistema linear se os parâmetros do sistema são modificados periodicamente com o tempo. Um sistema linear conservativo pode também apresentar oscilações que se parecem com oscilações sub-harmônicas de sistemas não lineares, mas as oscilações em sistemas lineares são essencialmente diferentes das oscilações sub-harmônicas.)

Oscilações auto-excitadas ou ciclos limite. Um outro fenômeno que é observado em certos sistemas não lineares é uma oscilação auto-excitada ou ciclo limite. Considere um sistema descrito pela seguinte equação:

$$m\ddot{x} - f(1 - x^2)\dot{x} + kx = 0$$

onde m, f e k são grandezas positivas. Esta equação é chamada equação de Van der Pol. Ela é não linear no termo de amortecimento. Ao examinar este termo, observamos que para pequenos valores de x o amortecimento será negativo e, na realidade, adicionará energia ao sistema, enquanto que para valores grandes de x ele será positivo, removendo energia do sistema. Portanto, pode-se esperar que tal sistema apresente uma oscilação mantida. Como este não é um sistema forçado, esta oscilação é chamada de oscilação auto-excitada ou ciclo limite. Note que se um sistema apresenta apenas um ciclo limite, como no caso do sistema presente, a amplitude deste ciclo limite não depende da condição inicial.

Sincronismo de freqüência. Um exemplo de um fenômeno interessante que pode ser observado em alguns sistemas não lineares é o sincronismo de freqüência. Se uma força periódica de freqüência ω é aplicada ao sistema capaz de apresentar um ciclo limite de freqüência ω_0, o fenômeno bem conhecido de batimento é observado. Quando a diferença entre as duas freqüências decresce, a freqüência do batimento também decresce. Em um sistema linear se determina, tanto experimental quanto teoricamente, que a freqüência de batimento decresce indefinidamente quando ω tende a ω_0. Entretanto, em um sistema não linear auto-excitado se determina experimentalmente que a freqüência ω_0 do ciclo limite entra em sincronismo, ou é sincronizada, pela freqüência de excitação ω dentro de uma certa faixa de freqüências. Este fenômeno é normalmente chamado de sincronismo de freqüência, e a banda de freqüência em que o sincronismo ocorre é chamada zona de sincronismo.

A Fig. 11.6 mostra a relação entre $|\omega - \omega_0|$ e ω. Para um sistema linear, a relação entre $|\omega - \omega_0|$ e ω seguiria as linhas tracejadas e $|\omega - \omega_0|$ seria zero apenas

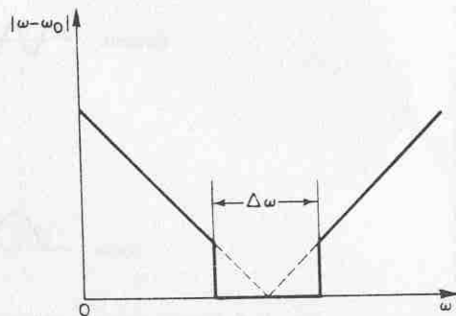


Fig. 11.6 Curva de $|\omega - \omega_0|$ em função de ω mostrando a região de sincronismo de freqüência.

para um valor de ω; $\omega = \omega_0$. Para um sistema não linear auto-excitado, ocorre o sincronismo de freqüência, e na zona de sincronismo, que é indicada pela região $\Delta\omega$ na Fig. 11.6, as freqüências ω e ω_0 se tornam iguais e existe apenas uma freqüência ω. Tal sincronismo de freqüência é observado na resposta de freqüência de sistemas não lineares que apresentam ciclos-limite.

Supressão assíncrona. Em um sistema não linear que exibe um ciclo limite de freqüência ω_0, é possível suprimir a oscilação de ciclo limite forçando o sistema a uma freqüência ω_1, onde ω_1 e ω_0 não são relacionadas uma com a outra. Este fenômeno é chamado de supressão assíncrona, ou estabilização de sinal.

Comentário. Nenhum dos fenômenos mencionados acima, bem como outros fenômenos não lineares não mencionados aqui, ocorre em sistemas lineares. Estes fenômenos não podem ser explicados pela teoria linear; para explicá-los, devemos resolver as equações diferenciais não lineares que descrevem a dinâmica do sistema em forma analítica ou gráfica.

11.2 SISTEMAS DE CONTROLE NÃO LINEARES

Muitos tipos diferentes de não linearidades podem ser achados em sistemas de controle reais e divididos em duas classes, dependendo de serem inerentes ao sistema ou inseridos propositadamente no sistema.

A seguir, primeiro discutiremos não linearidades inerentes, e então as não linearidades intencionais. Finalmente discutiremos abordagens à análise e projeto de sistemas de controle não lineares.

Não linearidades inerentes. As não linearidades inerentes são inevitáveis em sistemas de controle. Exemplos de tais não linearidades são

1. saturação
2. zona morta
3. histerese
4. folga (*backlash*)
5. atrito estático, atrito de Coulomb, e outros atritos não lineares
6. mola não linear
7. compressibilidade de fluido

De uma forma geral, a presença de tais não linearidades no sistema de controle afeta prejudicialmente o desempenho do sistema. Por exemplo, a folga pode causar instabilidade no sistema, e a zona morta pode causar erro estacionário.

Não linearidades intencionais. Alguns elementos não lineares são intencionalmente introduzidos em um sistema para melhorar o desempenho do sistema ou simplificar a construção deste, ou ambos. Um sistema não linear adequadamente projetado para uma certa função geralmente é superior do ponto de vista econômico, de peso, espaço, e confiabilidade quando comparado com sistemas lineares projetados para a mesma função. O exemplo mais simples de um sistema não linear intencional é um sistema operado convenientemente a relé. Outros exemplos podem ser achados em sistemas de controle ótimo que freqüentemente empregam controladores não lineares complicados. (Para uma discussão de sistemas de controle ótimo, referir-se ao Cap. 16.) Deve-se notar que embora elementos não lineares intencionais possam melhorar o desempenho do sistema sob certas condi-

ções especificadas de operação, em geral eles degradarão o desempenho do sistema sob outras condições de operação.

Efeito de não linearidades inerentes na precisão estática. Uma característica de sistemas de controle é que a potência é transmitida através do ramo direto, enquanto que a precisão estática do sistema é determinada pelos elementos no ramo de realimentação. Portanto, o dispositivo de medida determina o limite superior na precisão estática; a precisão estática não pode ser melhor que a precisão deste dispositivo de medida. Portanto, quaisquer não linearidades inerentes nos elementos de realimentação devem ser minimizadas.

Se os elementos de realimentação sofrem os efeitos de atrito, folga etc., então é desejável aplicar o sinal de erro em um dispositivo de integração, porque o sistema não pode detectar um erro muito pequeno, a menos que o erro pequeno seja integrado continuamente, tornando a amplitude grande suficiente para ser detectada.

Abordagens usadas na análise e projeto de sistemas de controle não lineares. Não há um método geral para se lidar com todos os sistemas não lineares porque as equações diferenciais não lineares praticamente não permitem um método geral de estudo. (Soluções exatas podem ser achadas apenas para certas equações diferenciais não lineares simples. Para muitas das equações diferenciais não lineares de importância prática, são possíveis apenas soluções aproximadas, e estas soluções se aplicam somente sob as condições limitadas em que foram obtidas.) Como não há um método geral, podemos tomar cada equação não linear, ou grupo de equações similares, individualmente e tentar desenvolver um método de análise que se aplicará satisfatoriamente para aquele grupo particular. (Note que embora se possa fazer uma quantidade muito pequena de generalizações dentro do grupo de equações similares, é impossível uma generalização ampla de uma solução particular.)

Uma maneira de se analisar e projetar um grupo particular de sistemas de controle não lineares, em que o grau da não linearidade é pequeno, é usar técnicas de linearização equivalente e resolver o problema linearizado obtido. O método da função descritiva a ser discutido neste capítulo é um dos métodos de linearização equivalente. Em muitos casos particulares, estamos primordialmente interessados na estabilidade de sistemas de controle não lineares, e soluções analíticas de equações diferenciais não lineares podem não ser necessárias. (O estabelecimento de critérios de estabilidade é muito mais fácil do que a obtenção de soluções analíticas.) O método da função descritiva nos permite estudar a estabilidade de muitos sistemas de controle não lineares simples do ponto de vista do domínio da freqüência.

O método da função descritiva fornece informação sobre a estabilidade para um sistema de qualquer ordem, mas não dá informação exata sobre as características de resposta temporal.

Outras formas de analisar e projetar sistemas de controle não lineares, que podem ser altamente não lineares, incluem a atual solução das equações diferenciais não lineares ou simplificações destas por meio de técnica do plano de fase ou aplicando técnicas baseadas no segundo método de Liapunov. (Para discussões sobre o método do plano de fase e do segundo método de Liapunov, referir-se aos Caps. 12 e 15, respectivamente.)

O método do plano de fase fornece informação tanto sobre estabilidade quando comportamento de resposta temporal, mas é limitado a sistemas de primeira e segunda-ordem.

O segundo método de Liapunov pode ser aplicado na análise de estabilidade de qualquer sistema não linear, mas sua aplicação pode ser prejudicada devido à

dificuldade de se achar funções de Liapunov para sistemas não lineares complicados. (Para funções de Liapunov, vide Cap. 15.)

Soluções em computador de problemas não lineares. Computadores modernos levaram a métodos novos de se estudar problemas não lineares. Técnicas de simulação em computadores pelo uso de computadores analógicos e/ou digitais são muito poderosas para analisar e projetar sistemas de controle não lineares. Agora é possível estudar sistemas não lineares bastante complicados através do uso de computadores em um espaço de tempo pequeno. Quando a complexidade de um sistema não permite o uso da abordagem analítica, as simulações em computadores podem ser muito vantajosas para se obter a informação necessária para fins de projeto.

Comentário. É importante lembrar que embora a predição do comportamento de sistemas não lineares seja normalmente difícil, ao se projetar um sistema de controle não devemos tentar forçar o sistema a ser o mais linear possível, porque a exigência de linearidade do sistema pode levar ao projeto de um sistema caro e menos desejável do que um sistema não linear adequadamente projetado.

11.3 FUNÇÕES DESCRITIVAS

Esta seção apresenta representações de função descritiva de elementos não lineares comumente encontrados.

Funções descritivas. Suponha que a entrada para um elemento não linear é senoidal. A saída do elemento não linear é, em geral, não senoidal. Suponha que a saída é periódica com o mesmo período que a entrada. (A saída contém harmônicas superiores, em adição à componente harmônica fundamental.)

Na análise por função descritiva, supomos que apenas a componente harmônica fundamental da saída é significativa. Tal suposição é freqüentemente válida uma vez que harmônicas superiores na saída de um elemento não linear são freqüentemente de menor amplitude do que a amplitude da componente de harmônica fundamental. Além disso, a maioria de sistemas de controle são filtros passa-baixas, com o resultado de que as harmônicas superiores são muito atenuadas quando comparadas com a componente harmônica fundamental.

A função descritiva ou função descritiva senoidal de um elemento não linear é definida como a relação complexa entre a componente harmônica fundamental da saída e a amplitude da entrada. Isto é,

$$N = \frac{Y_1}{X} \underline{/\phi_1}$$

onde

N = função descritiva
X = amplitude da senóide de entrada
Y_1 = amplitude da componente harmônica fundamental da saída
ϕ_1 = defasagem da componente harmônica fundamental da saída

Se não há elemento de armazenamento de energia incluído no elemento não linear, então N é uma função apenas da amplitude da entrada para o elemento. Por outro lado, se elementos armazenadores de energia são incluídos, então N é uma função tanto da amplitude como da freqüência da entrada.

Ao se calcular a função descritiva para um dado elemento não linear, necessi-

tamos achar a componente harmônica fundamental da saída. Para a entrada senoidal $x(t) = X \operatorname{sen} \omega t$ para o elemento não linear, a saída $y(t)$ pode ser expressa como uma série de Fourier como segue:

$$y(t) = A_0 + \sum_{n=1}^{\infty} (A_n \cos n\omega t + B_n \operatorname{sen} n\omega t) = A_0 + \sum_{n=1}^{\infty} Y_n \operatorname{sen}(n\omega t + \phi_n)$$

onde

$$A_n = \frac{1}{\pi}\int_0^{2\pi} y(t) \cos n\omega t \, d(\omega t)$$

$$B_n = \frac{1}{\pi}\int_0^{2\pi} y(t) \operatorname{sen} n\omega t \, d(\omega t)$$

$$Y_n = \sqrt{A_n^2 + B_n^2}$$

$$\phi_n = \tan^{-1}\left(\frac{A_n}{B_n}\right)$$

Tabela 11.1 Três não linearidades e suas funções descritivas

y, k_2, k_1, S, x	$N = k_2 + \dfrac{2(k_1 - k_2)}{\pi}\left(\operatorname{sen}^{-1}\dfrac{S}{X} + \dfrac{S}{X}\sqrt{1 - \dfrac{S^2}{X^2}}\right)$ $(X \geq S)$
y, k, M, x	$N = k + \dfrac{4M}{\pi X}$
y, k, Δ, Δ	$N = k - \dfrac{2k}{\pi}\operatorname{sen}^{-1}\dfrac{\Delta}{X} + \dfrac{(4 - 2k)\Delta}{\pi X}\sqrt{1 - \dfrac{\Delta^2}{X^2}}$ $(X \geq \Delta)$

Se a não linearidade é simétrica, então $A_0 = 0$. A componente da harmônica fundamental da saída é

$$y_1(t) = A_1 \cos \omega t + B_1 \operatorname{sen} \omega t$$
$$= Y_1 \operatorname{sen}(\omega t + \phi_1)$$

A função descritiva é então dada por

$$N = \frac{Y_1}{X} \angle \phi_1 = \frac{\sqrt{A_1^2 + B_1^2}}{X} \angle \tan^{-1}\left(\frac{A_1}{B_1}\right)$$

Claramente, N é uma grandeza complexa quando ϕ_1 é não nula.

A Tabela 11.1 mostra três não linearidades e suas funções descritivas. (Na Tabela 11.1, k_1, k_2 e k indicam as inclinações das linhas.) Cálculos ilustrativos das funções descritivas de não linearidades comumente encontradas são dados a seguir.

Não linearidade tipo liga-desliga (*on-off*). A não linearidade tipo liga-desliga é muitas vezes chamada de uma não linearidade de duas posições. Considere um elemento tipo liga-desliga cuja curva característica entrada-saída é vista na Fig. 11.7(a). A saída deste elemento é ou uma constante positiva ou uma constante negativa, e a Fig. 11.7(b) mostra as formas de onda da entrada e da saída.

Vamos obter uma expansão em série de Fourier da saída $y(t)$ de tal elemento.

$$y(t) = A_0 + \sum_{n=1}^{\infty} (A_n \cos n\omega t + B_n \operatorname{sen} n\omega t)$$

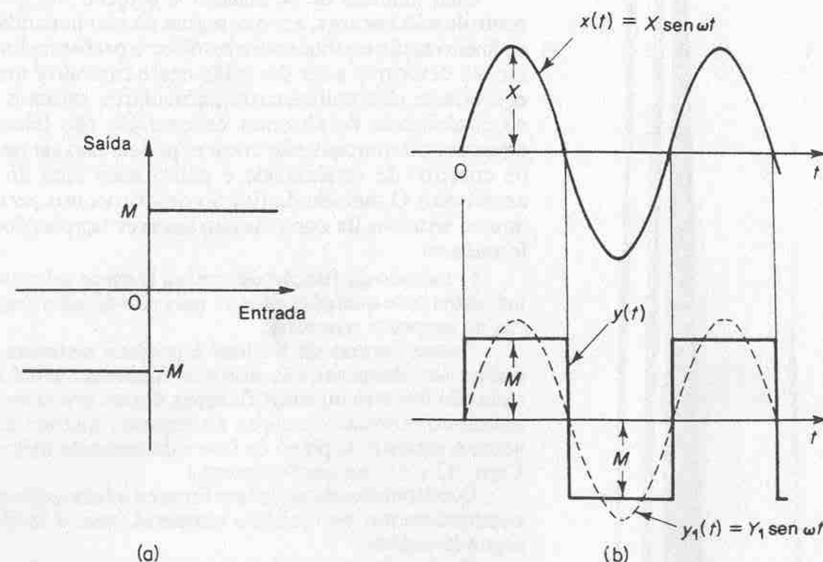


Fig. 11.7 (a) Curva da característica de entrada-saída para a não linearidade tipo liga-desliga; (b) formas de onda de entrada e saída para a não linearidade tipo liga-desliga.

Como visto na Fig. 11.7(b), a saída é uma função ímpar. Para uma função ímpar, temos $A_n = 0$ $(n = 0, 1, 2, \ldots)$. Portanto.

$$y(t) = \sum_{n=1}^{\infty} B_n \operatorname{sen} n\omega t$$

A componente harmônica fundamental de $y(t)$ é

$$y_1(t) = B_1 \operatorname{sen} \omega t = Y_1 \operatorname{sen} \omega t$$

onde

$$Y_1 = \frac{1}{\pi} \int_0^{2\pi} y(t) \operatorname{sen} \omega t \, d(\omega t) = \frac{2}{\pi} \int_0^{\pi} y(t) \operatorname{sen} \omega t \, d(\omega t)$$

Substituindo $y(t) = M$ nesta última equação, resulta

$$Y_1 = \frac{2M}{\pi} \int_0^{\pi} \operatorname{sen} \omega t \, d(\omega t) = \frac{4M}{\pi}$$

Portanto,

$$y_1(t) = \frac{4M}{\pi} \operatorname{sen} \omega t$$

A função descritiva N é então dada por

$$N = \frac{Y_1}{X} \underline{/0^\circ} = \frac{4M}{\pi X}$$

Claramente, a função descritiva de um elemento tipo liga-desliga é uma grandeza real e uma função apenas da amplitude de entrada X. Um gráfico desta função descritiva em função de M/X é visto na Fig. 11.8.

Não linearidade tipo liga-desliga com histerese. Considere um elemento tipo liga-desliga com histerese cuja curva característica de entrada-saída é vista na Fig. 11.9(a). As formas de onda da entrada e saída são vistas na Fig. 11.9(b). Claramente, a saída é uma onda quadrada, mas está atrasada em relação à entrada por $\omega t_1 = \operatorname{sen}^{-1}(h/X)$. Portanto a função descritiva para este elemento não linear é

$$N = \frac{4M}{\pi X} \underline{\Big/ -\operatorname{sen}^{-1}\left(\frac{h}{X}\right)}$$

É conveniente fazer um gráfico de

$$\frac{h}{M} N = \frac{4h}{\pi X} \underline{\Big/ -\operatorname{sen}^{-1}\left(\frac{h}{X}\right)}$$

em função de h/X ao invés de N em função de h/X porque hN/M é uma função apenas de h/X. Um gráfico de hN/M em função de h/X é visto na Fig. 11.10.

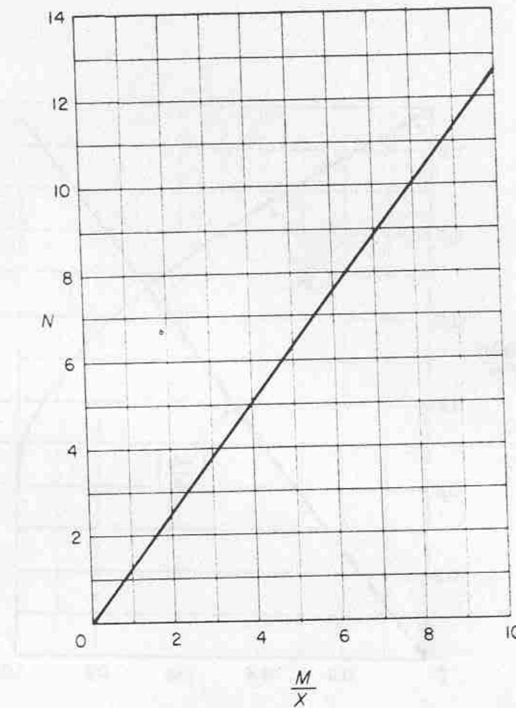


Fig. 11.8 Função descritiva da não linearidade tipo liga-desliga.

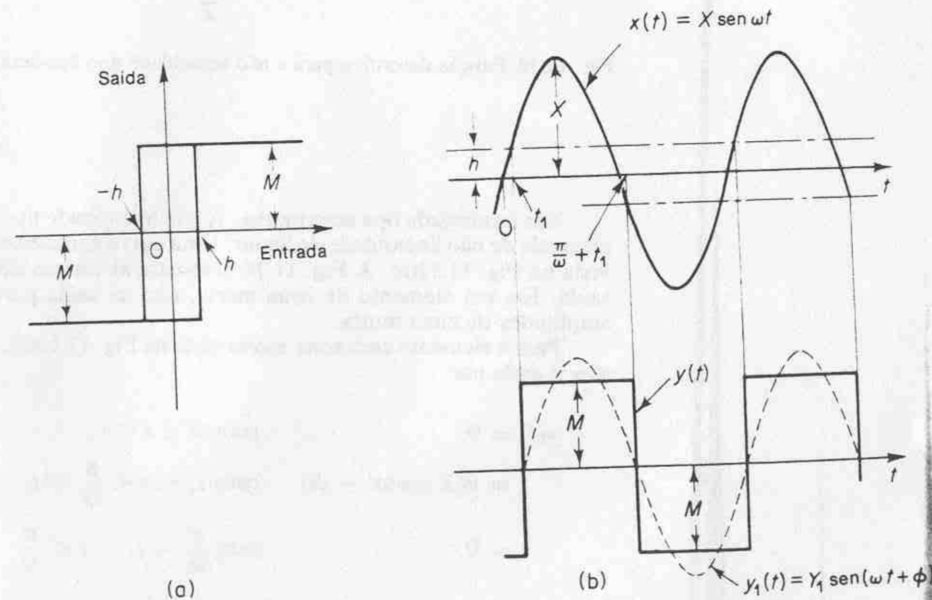


Fig. 11.9 (a) Curva da característica de entrada-saída para a não linearidade tipo liga-desliga com histerese; (b) formas de onda de entrada e saída para a não linearidade tipo liga-desliga com histerese.

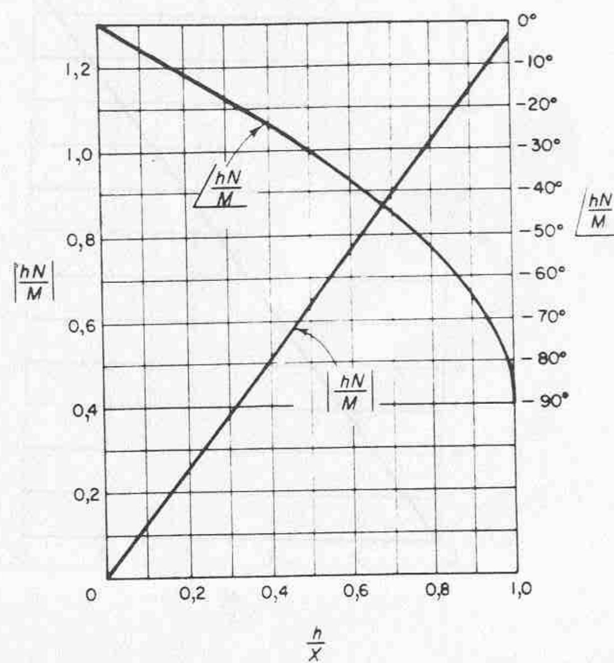


Fig. 11.10 Função descritiva para a não linearidade tipo liga-desliga com histerese.

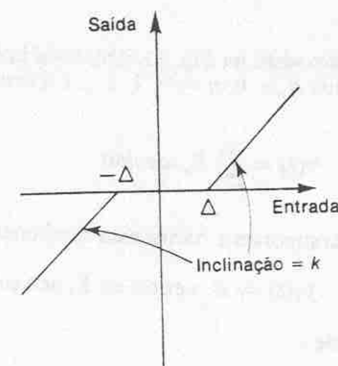
Não linearidade tipo zona morta. A não linearidade tipo zona morta é às vezes chamada de não linearidade de limiar. Uma curva característica de entrada-saída é vista na Fig. 11.11(a). A Fig. 11.11(b) mostra as formas de onda da entrada e da saída. Em um elemento de zona morta, não há saída para entradas dentro das amplitudes de zona morta.

Para o elemento com zona morta visto na Fig. 11.11(a), a saída $y(t)$ para $0 \leq \omega t \leq \pi$ é dada por

$$
\begin{aligned}
y(t) &= 0 && \text{para } 0 < t < t_1 \\
&= k(X \operatorname{sen}\omega t - \Delta) && \text{para } t_1 < t < \frac{\pi}{\omega} - t_1 \\
&= 0 && \text{para } \frac{\pi}{\omega} - t_1 < t < \frac{\pi}{\omega}
\end{aligned}
$$

Como a saída $y(t)$ é novamente uma função ímpar, sua expansão em série de Fourier tem apenas elementos senoidais. A componente harmônica fundamental da saída é dada por

$$y_1(t) = Y_1 \operatorname{sen}\omega t$$



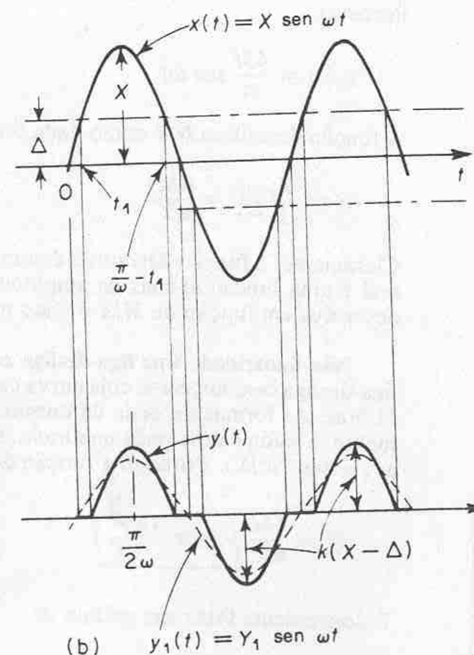


Fig. 11.11 (a) Curva característica de entrada-saída para a não linearidade tipo zona morta; (b) formas de onda da entrada e saída para a não linearidade tipo zona morta.

onde

$$Y_1 = \frac{1}{\pi}\int_0^{2\pi} y(t)\,\text{sen}\,\omega t\; d(\omega t)$$
$$= \frac{4}{\pi}\int_0^{\pi/2} y(t)\,\text{sen}\,\omega t\; d(\omega t)$$
$$= \frac{4k}{\pi}\int_{\omega t_1}^{\pi/2} (X\,\text{sen}\,\omega t - \Delta)\,\text{sen}\,\omega t\; d(\omega t)$$

Note que

$$\Delta = X\,\text{sen}\,\omega t_1$$

ou

$$\omega t_1 = \text{sen}^{-1}\left(\frac{\Delta}{X}\right)$$

Portanto,

$$Y_1 = \frac{4Xk}{\pi}\left[\int_{\omega t_1}^{\pi/2} \text{sen}^2\,\omega t\; d(\omega t) - \text{sen}\,\omega t_1 \int_{\omega t_1}^{\pi/2} \text{sen}\,\omega t\; d(\omega t)\right]$$
$$= \frac{2Xk}{\pi}\left[\frac{\pi}{2} - \text{sen}^{-1}\left(\frac{\Delta}{X}\right) - \frac{\Delta}{X}\sqrt{1-\left(\frac{\Delta}{X}\right)^2}\right]$$

A função descritiva para um elemento com zona morta pode ser obtida como

$$N = \frac{Y_1}{X}\underline{/0^\circ} = k - \frac{2k}{\pi}\left[\text{sen}^{-1}\frac{\Delta}{X} + \frac{\Delta}{X}\sqrt{1-\left(\frac{\Delta}{X}\right)^2}\right]$$

A Fig. 11.12 mostra um gráfico de N/k em função de Δ/X. Note que para $(\Delta/X) > 1$ a saída é nula, assim como o valor da função descritiva.

Não linearidade tipo saturação. Uma curva característica da entrada-saída para a não linearidade tipo saturação é vista na Fig. 11.13(a). Para sinais de entrada pequenos, a saída de um elemento de saturação é proporcional à entrada. Para sinais de entrada maiores a saída não aumentará proporcionalmente, e, finalmente, para sinais de entrada muito grandes a saída é constante. A Fig. 11.13(b) mostra as formas de onda da entrada e da saída para a não linearidade tipo saturação.

A função descritiva para um tal elemento pode ser obtida como

$$N = \frac{2k}{\pi}\left[\text{sen}^{-1}\left(\frac{S}{X}\right) + \frac{S}{X}\sqrt{1-\left(\frac{S}{X}\right)^2}\right]$$

A Fig. 11.14 mostra um gráfico de N/k em função de S/X. Para $(S/X) > 1$, o valor da função descritiva é unitário.

Observe que a função descritiva para a não linearidade tipo zona morta e aquela para a não linearidade tipo saturação são relacionadas como segue:

$$N_{\text{zona morta}} = k - N_{\text{saturação}} \qquad \text{para } \Delta = S$$

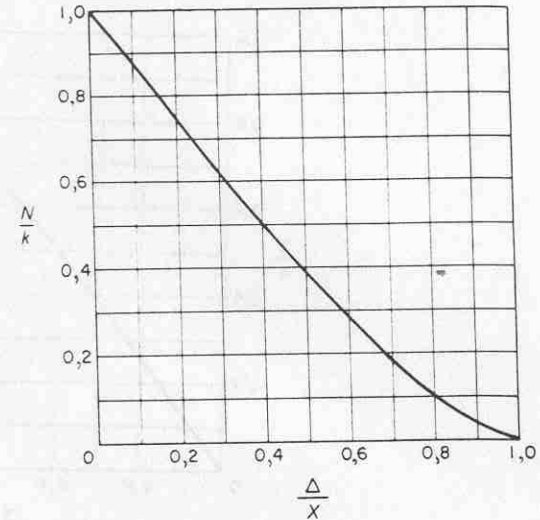


Fig. 11.12 Função descritiva para a não linearidade tipo zona morta.

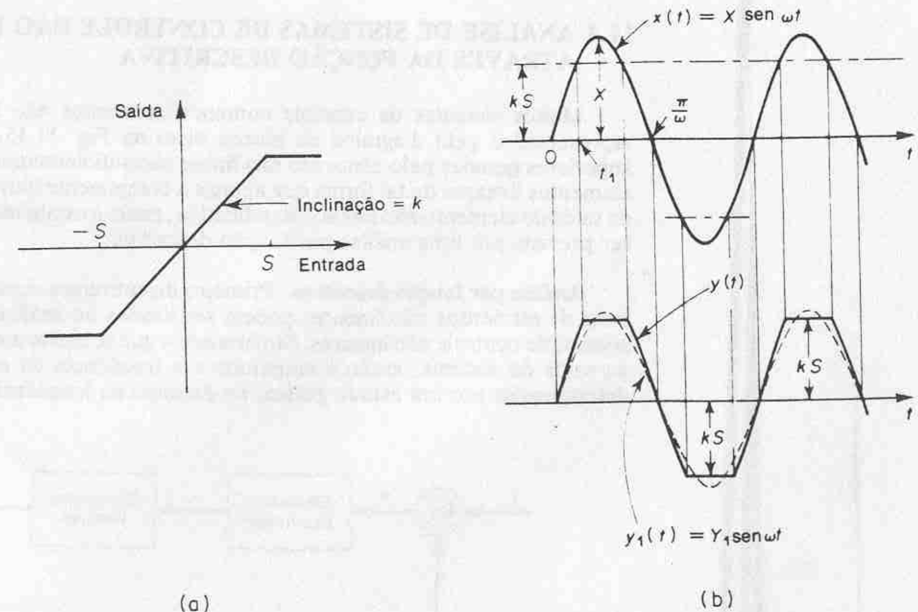


Fig. 11.13 (a) Curva de característica entrada-saída para a não linearidade tipo saturação; (b) formas de onda da entrada e saída para a não linearidade tipo saturação.

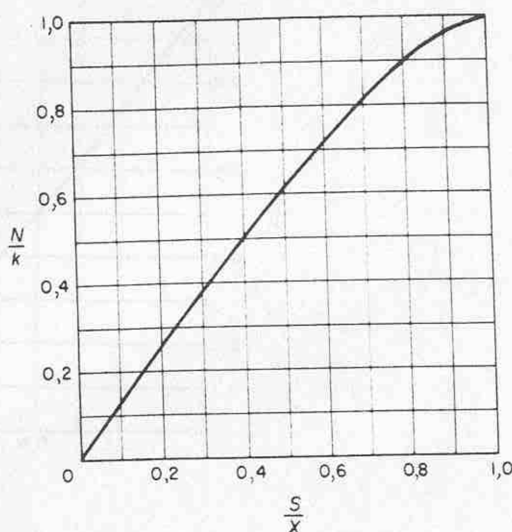


Fig. 11.14 Função descritiva para a não linearidade tipo saturação.

11.4 ANÁLISE DE SISTEMAS DE CONTROLE NÃO LINEARES ATRAVÉS DA FUNÇÃO DESCRITIVA

Muitos sistemas de controle contendo elementos não lineares podem ser representados pelo diagrama de blocos visto na Fig. 11.15. Se as harmônicas superiores geradas pelo elemento não linear são suficientemente atenuadas pelos elementos lineares de tal forma que apenas a componente harmônica fundamental da saída do elemento não linear é significativa, então a estabilidade do sistema pode ser prevista por uma análise por função descritiva.

Análise por função descritiva. Primeiro discutiremos como as funções descritivas de elementos não lineares podem ser usadas na análise de estabilidade de sistema de controle não lineares. Mostraremos que se existe uma oscilação mantida na saída do sistema, então a amplitude e a freqüência da oscilação podem ser determinadas por um estudo gráfico no domínio da freqüência.

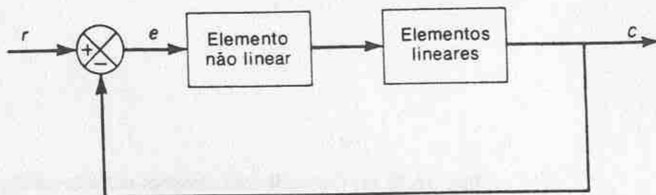


Fig. 11.15 Sistema de controle não linear.

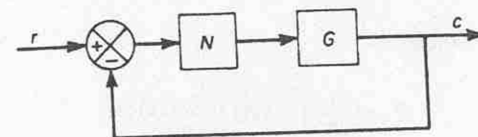


Fig. 11.16 Sistema de controle não linear.

Considere o sistema visto na Fig. 11.16. O bloco N indica a função descritiva do elemento não linear. Se as harmônicas de ordem superior são suficientemente atenuadas, a função descritiva N pode ser tratada como um ganho de variável real ou complexa. Então a resposta em freqüência de malha-fechada se torna

$$\frac{C(j\omega)}{R(j\omega)} = \frac{NG(j\omega)}{1 + NG(j\omega)}$$

A equação característica é

$$1 + NG(j\omega) = 0$$

ou

$$G(j\omega) = -\frac{1}{N} \tag{11.2}$$

Se a Eq. (11.2) é satisfeita, então a saída do sistema exibirá um ciclo limite. Esta situação corresponde ao caso em que o lugar geométrico de $G(j\omega)$ passa pelo ponto crítico. (Na análise convencional de resposta em freqüência de sistemas de controle lineares, o ponto crítico é o ponto $-1 + j0$.)

Na análise de função descritiva, a análise convencional da resposta em freqüência é modificada de tal forma que todo o lugar geométrico $-1/N$ se torna o lugar geométrico de pontos críticos. Portanto, a posição relativa dos lugares geométricos de $-1/N$ e de $G(j\omega)$ proverá a informação de estabilidade.

Para determinar a estabilidade do sistema, fazemos um gráfico dos lugares geométricos de $-1/N$ e de $G(j\omega)$. Na presente análise, supomos que a parte linear do sistema é de mínima fase ou que todos os pólos e zeros de $G(s)$ estão no semiplano esquerdo do plano s, incluindo o eixo $j\omega$. O critério para estabilidade é que se o lugar geométrico de $-1/N$ não é interceptado pelo lugar geométrico de $G(j\omega)$, então o sistema é estável, ou não há ciclo limite em regime estacionário.

Por outro lado, se o lugar geométrico de $-1/N$ é interceptado pelo lugar geométrico de $G(j\omega)$, então o sistema é instável, e a saída do sistema, quando sujeita a qualquer perturbação, crescerá até que haja danificação ou que se atinja um valor limite determinado por um obstáculo mecânico ou outro dispositivo de segurança.

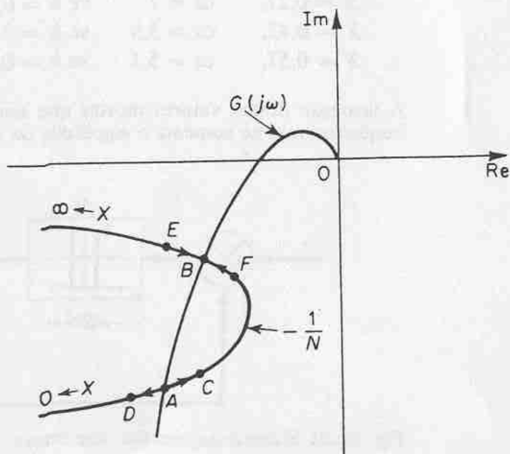
Se os lugares geométricos de $-1/N$ e de $G(j\omega)$ se interceptam, então a saída do sistema pode apresentar uma oscilação mantida, ou um ciclo limite. Tal oscilação mantida não é senoidal, mas pode ser aproximada por uma senóide. A oscilação mantida é caracterizada pelo valor de X no lugar geométrico de $-1/N$ e o valor de ω na intersecção do lugar geométrico de $G(j\omega)$.

Em geral, um sistema de controle não deve apresentar comportamento de ciclo limite, embora um ciclo limite de amplitude pequena possa ser aceitável em certas aplicações.

Estabilidade de oscilações mantidas, ou ciclos limite. A estabilidade de um ciclo limite pode ser prevista como segue: Considere o sistema visto na Fig. 11.17. Suponha que o ponto A no lugar geométrico de $-1/N$ corresponde a um baixo valor de X, onde X é a amplitude do sinal de entrada senoidal no elemento não linear, e que o ponto B no lugar geométrico de $-1/N$ corresponde a um valor grande de X. O valor de X no lugar geométrico de $-1/N$ aumenta na direção do ponto A ao ponto B.

Vamos supor que o sistema está sendo operado no ponto A. A oscilação tem amplitude X_A e freqüência ω_A, determinados a partir dos lugares geométricos de $-1/N$ e de $G(j\omega)$, respectivamente. Suponha que uma leve perturbação é dada ao sistema, sendo operado no ponto A de tal forma que a amplitude da entrada ao elemento não linear é levemente aumentada. (Por exemplo, suponha que o ponto de operação se move do ponto A ao ponto C no lugar geométrico de $-1/N$.) Então o ponto de operação C corresponde ao ponto crítico ou ao ponto $-1 + j0$ no plano complexo para sistemas de controle lineares. Portanto, como visto na Fig. 11.17, o lugar geométrico de $G(j\omega)$ circunda o ponto C no significado de Nyquist. Como isto é semelhante ao caso em que o lugar geométrico de malha-aberta de um sistema linear circunda o ponto $-1 + j0$, a amplitude aumentará e o ponto de operação se move em direção ao ponto B.

Em seguida suponha que uma pequena perturbação diminui a amplitude da entrada senoidal ao elemento não linear. Suponha que o ponto de operação é movido do ponto A ao ponto D no lugar geométrico de $-1/N$. O ponto D então corresponde ao ponto crítico. Neste caso, o lugar geométrico de $G(j\omega)$ não circunda o ponto crítico e, portanto, a amplitude da entrada do elemento não linear diminui, e



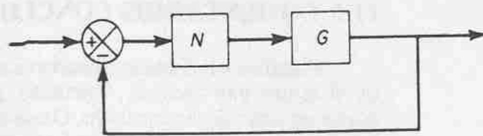


Fig. 11.17 Análise de estabilidade de operações de ciclo limite de sistema de controle não linear.

o ponto de operação se move além do ponto D para a esquerda. Portanto, o ponto A possui características divergentes e corresponde a um ciclo limite instável.

Considere a seguir o caso em que uma pequena perturbação é dada ao sistema operando no ponto B. Suponha que o ponto de operação é movido para o ponto E no lugar geométrico de $-1/N$. Então, neste caso, o lugar geométrico de $G(j\omega)$ não circunda o ponto crítico (ponto E). A amplitude da senóide de entrada do elemento não linear decresce, e o ponto de operação se move em direção ao ponto B.

De forma similar, suponha que uma pequena perturbação faz com que o ponto de operação do sistema se mova do ponto B ao ponto F. Então o lugar geométrico de $G(j\omega)$ circundará o ponto crítico (ponto F). Portanto, a amplitude da oscilação crescerá, e o ponto de operação se move do ponto F em direção ao ponto B. Desta forma, o ponto B possui características convergentes, e a operação do sistema no ponto B é estável; em outras palavras, o ciclo limite neste ponto é estável.

Para o sistema visto na Fig. 11.17, o ciclo limite estável correspondendo ao ponto B pode ser observado experimentalmente, mas não o ciclo limite instável que corresponde ao ponto A.

Precisão da análise por função descritiva. Observe que a amplitude e freqüência do ciclo limite indicado pela intersecção dos lugares geométricos de $-1/N$ e $G(j\omega)$ são valores aproximados.

Se os lugares geométricos de $-1/N$ e $G(j\omega)$ se interceptam quase perpendicularmente, então a precisão da análise por função descritiva é geralmente boa. (Se as harmônicas de ordem superior são todas atenuadas, a precisão é excelente. Caso contrário a precisão é boa a razoável.)

Se o lugar geométrico de $G(j\omega)$ é tangente, ou quase tangente ao lugar geométrico de $-1/N$, então a precisão da informação obtida da análise por função descritiva depende de quão bem irá $G(j\omega)$ atenuar as harmônicas de ordem superior. Em alguns casos, há uma oscilação mantida, em outros casos, não há tais oscilações. Este fato depende da natureza de $G(j\omega)$. Entretanto, pode-se dizer que o sistema está a ponto de apresentar um ciclo limite quando os lugares geométricos de $-1/N$ e de $G(j\omega)$ são tangentes entre si.

Exemplo 11.1 A Fig. 11.18 mostra um sistema de controle com uma não linearidade tipo saturação. Supomos que $G(s)$ é uma função de transferência de mínima fase. A Fig. 11.19 mostra um gráfico dos lugares geométricos de $-1/N$ e $G(j\omega)$. O lugar geométrico de $-1/N$ começa no ponto -1 no eixo real negativo e se estende até $-\infty$. Claramente N é uma função apenas da amplitude do sinal de entrada $x(t) = X \text{ sen } \omega t$. O lugar geométrico de $G(j\omega)$ é uma função apenas de ω.

A Fig. 11.19 mostra que os dois lugares geométricos se interceptam. Esta intersecção corresponde a um ciclo limite estável. A amplitude do ciclo limite é lida do lugar geométrico de $-1/N$ como sendo $X = X_1$. A freqüência do ciclo limite é lida do lugar geométrico de $G(j\omega)$ como sendo $\omega = \omega_1$.

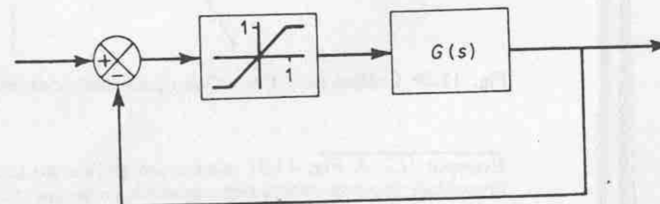


Fig. 11.18 Sistema de controle com não linearidade tipo saturação.

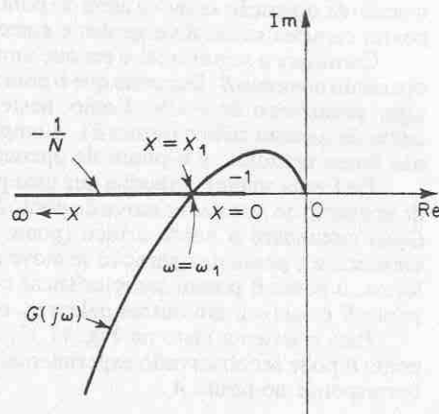


Fig. 11.19 Gráfico de $-1/N$ e $G(j\omega)$ para análise de estabilidade.

Na ausência de qualquer entrada de referência, a saída deste sistema em regime estacionário apresenta uma oscilação mantida com amplitude igual a X_1 e freqüência igual a ω_1.

Se o ganho da função de transferência $G(s)$ é diminuído de tal forma que os lugares geométricos de $-1/N$ e $G(j\omega)$ não se interceptem, como visto na Fig. 11.20, então o sistema se torna estável, e quaisquer oscilações que possam ocorrer na saída do sistema como resultado de perturbações se extinguirão e não haverá oscilações mantidas em regime estacionário. Isto acontece porque o lugar geométrico de $-1/N$ está à esquerda do lugar geométrico de $G(j\omega)$, ou o lugar geométrico de $G(j\omega)$ não intercepta o lugar geométrico de $-1/N$.

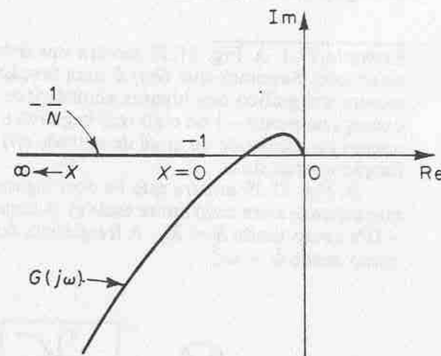


Fig. 11.20 Gráfico de $-1/N$ e $G(j\omega)$ para análise de estabilidade.

Exemplo 11.2 A Fig. 11.21 mostra um gráfico do lugar geométrico de $-1/N$ para a não linearidade tipo zona morta bem como o lugar geométrico de $G(j\omega)$. Neste sistema, os lugares geométricos de $-1/N$ e $G(j\omega)$ se interceptam. O ciclo limite neste caso é instável. A oscilação ou se extingue ou aumenta de amplitude indefinidamente. Isto indica uma situação indesejável e deve ser evitada.

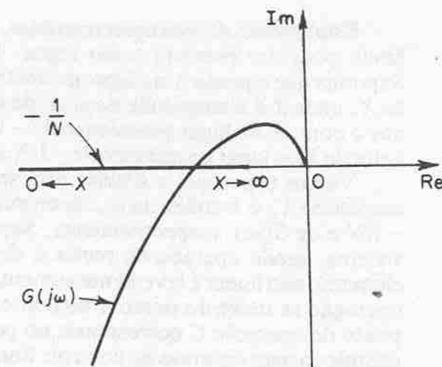


Fig. 11.21 Gráfico de $-1/N$ e $G(j\omega)$ para análise de estabilidade.

Exemplo 11.3 Considere o sistema visto na Fig. 11.22. Determine o efeito de histerese na amplitude e freqüência da operação de ciclo limite do sistema.

Os lugares geométricos de $-1/N$ para três diferentes valores de h : $h = 0,1$, $0,2$ e $0,3$, são vistos na Fig. 11.23, juntamente com o lugar geométrico de $G(j\omega)$. Os lugares geométricos de $-1/N$ são retas paralelas ao eixo real. Os valores de N são obtidos a partir da Fig. 11.10.

Da Fig. 11.23, pode-se ver que as amplitudes e freqüências dos ciclos limite são

$$X = 0{,}27, \quad \omega = 7 \quad \text{se } h = 0{,}1$$
$$X = 0{,}42, \quad \omega = 5{,}9 \quad \text{se } h = 0{,}2$$
$$X = 0{,}57, \quad \omega = 5{,}1 \quad \text{se } h = 0{,}3$$

A inspeção destes valores mostra que aumentando a largura da histerese se diminui a freqüência mas se aumenta a amplitude do ciclo limite, como esperado.

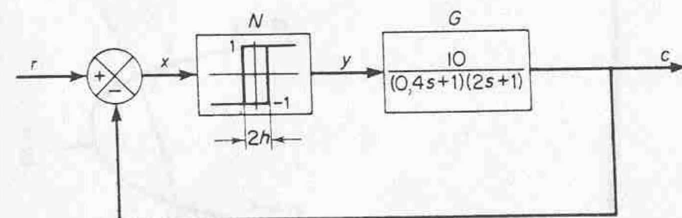


Fig. 11.22 Sistema de controle não linear.

11.5 COMENTÁRIOS CONCLUSIVOS

A análise por função descritiva é uma extensão de técnicas lineares ao estudo de sistemas não lineares. Portanto, aplicações típicas são para sistemas com um baixo grau de não linearidade. O uso da função descritiva na análise de sistemas não lineares de grau elevado pode levar a resultados muito errados; isto limita a aplicabilidade da função descritiva para a análise e projeto de sistemas não lineares de baixo grau.

Para concluir este capítulo, resumiremos o método da função descritiva para a análise e projeto de sistemas de controle não lineares.

1. O método da função descritiva é um método aproximado para determinar a estabilidade de sistemas de controle não lineares não excitados. Ao aplicar este método, deve-se ter em mente as suposições básicas e limitações. Embora muitos sistemas de controle na prática satisfaçam as suposições básicas do método de função descritiva, há alguns sistemas que não o fazem. Portanto, sempre é necessário examinar a validade do método em cada caso.

2. Na análise por função descritiva, a natureza da não linearidade presente no sistema determina a complexidade da análise. Em outras palavras, os elementos lineares de qualquer ordem não afetam a complexidade. Uma vantagem deste método é que a análise não é essencialmente mais complicada para sistemas com dinâmica complexa nas suas partes lineares. A precisão da análise é melhor para sistemas de ordem alta do que de ordem baixa porque sistemas de ordem elevada em geral têm melhores características de filtragem passa-baixas.

3. Embora o método da função descritiva seja bastante útil na predição da estabilidade de sistemas não forçados, ele fornece pouca informação sobre características de resposta transitória.

4. O método da função descritiva é conveniente para ser aplicado a problemas de projeto. O uso de funções descritivas nos permite aplicar métodos de resposta em freqüência para o reajuste ou modificação do lugar geométrico de $G(j\omega)$. A análise por função descritiva é particularmente útil quando o projetista quer ter uma idéia a grosso modo dos efeitos de certas não linearidades ou dos efeitos da modificação de componentes lineares ou não lineares na malha de controle. A análise fornece informação gráfica sobre a estabilidade e sugere maneiras de melhorar as características de resposta, caso necessário. Quando o lugar geométrico de $-1/N$ e de $G(j\omega)$ são colocados em um

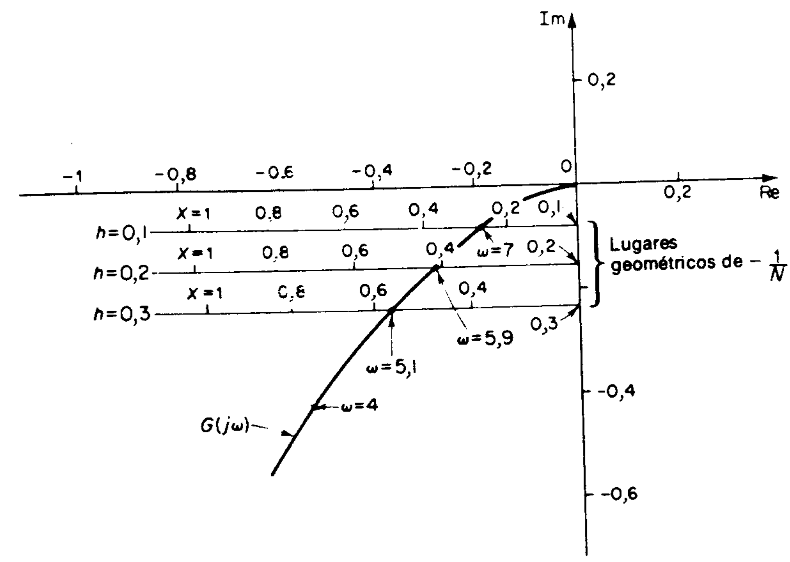


Fig. 11.23 Gráfico de $-1/N$ e $G(j\omega)$ para o sistema visto na Fig. 11.22.

gráfico do plano complexo, o desempenho do sistema pode ser estimado rapidamente a partir do gráfico. Se quaisquer melhoramentos de desempenho são necessários, eles podem ser conseguidos modificando-se os lugares geométricos. Esta modificação dos lugares geométricos sugere o tipo de uma rede de compensação adequada. (Projetos através da função descritiva podem sugerir certo lugar geométrico de $-1/N$, em particular, ao invés de sugerir a modificação do lugar geométrico de $G(j\omega)$. Entretanto, a realização física de um elemento não linear com uma função descritiva especificada pode ser difícil.) É interessante notar que embora o método da função descritiva nos permita predizer ciclos-limite com boa precisão em termos de engenharia, em problemas de projeto o método é usado como um critério negativo, em que parâmetros do sistema são ajustados até que condições de ciclo limite sejam eliminadas e garantida uma estabilidade relativa adequada.

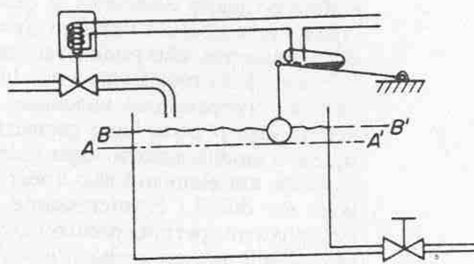
5. Um sistema físico pode possuir dois ou mais elementos não lineares. Quando apenas um elemento não linear se torna significativo para uma particular condição de operação, os efeitos dos outros elementos não lineares podem ser desprezados na análise. Por exemplo, se o sistema apresenta tanto não linearidade para baixos sinais como não linearidades para sinais grandes, o primeiro pode ser desprezado quando a amplitude do sinal é grande e vice-versa. É importante manter em mente que a função descritiva de duas não linearidades em série é, em geral, diferente do produto das funções descritivas individuais. Portanto, se dois ou mais elementos não lineares, que não são separados entre si por filtros passa-baixas efetivos, se tornam simultaneamente significativos sob certas condições de operação, eles podem ser combinados em um bloco, e a função descritiva equivalente deste bloco pode ser determinada. Neste caso, a função descritiva pode-se tornar tanto dependente da amplitude como da freqüência.

6. Na análise por função descritiva usual, supõe-se que a entrada para a não linearidade é senoidal, mas esta hipótese pode ser estendida. A entrada para a não linearidade pode ser um sinal senoidal, somado com um sinal adicional, embora esta complicação adicional possa tornar a análise muito maçante. As funções descritivas que correspondem a este caso são chamadas de funções descritivas de entradas duais.

7. Em alguns casos a análise de estabilidade de sistemas de controle pode ser de interesse primordial, mas em outros casos pode-se desejar uma resposta ótima (em certo sentido). Problemas de projeto ótimo podem envolver a determinação de um controlador não linear (ou computador) a ser inserido no sistema. O desempenho do sistema não linear depende bastante dos sinais de entrada. Isto significa que são necessárias descrições precisas da entrada e da saída desejada. Em virtude da pequena correlação entre a resposta em freqüência e a resposta temporal de sistemas não lineares, o método da função descritiva deixa de ser útil no projeto de sistemas de controle ótimos com entradas aperiódicas. (Sistemas de controle ótimos com entradas aperiódicas são discutidos no Cap. 16.)

Problemas Ilustrativos e Soluções

Problema A.11.1 A Fig. 11.24(a) é um diagrama esquemático de um sistema de controle de nível de líquido. O movimento da bóia posiciona a chave elétrica de mercúrio que energiza ou desenergiza a válvula elétrica operada por solenóide. Quando a válvula está aberta, entra líquido no tanque. A ação de controle é liga-desliga com histerese. A curva da taxa de influxo em função do erro é vista na Fig. 11.24(b). (A largura da histerese $2h$ é chamada de abertura *(gap)* diferencial.)



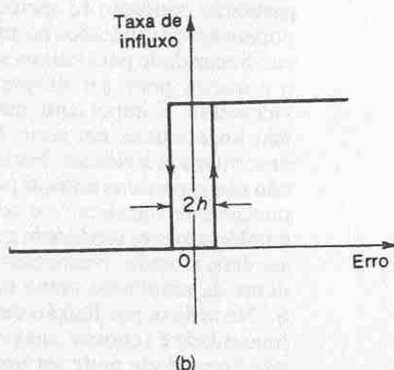


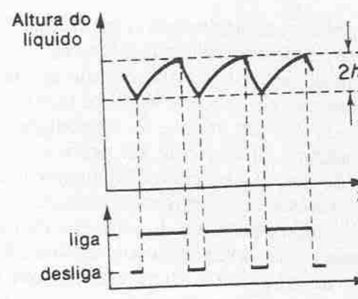
Fig. 11.24 (a) Sistema de controle de nível de líquido; (b) curva da taxa de influxo em função do erro para o controlador.

Faça um gráfico da altura do líquido em função do tempo em operação estacionária sob as seguintes condições:

1. A taxa de subida na altura do líquido quando a válvula de influxo está aberta é consideravelmente menor que a taxa de descida na altura do líquido quando a válvula está fechada.
2. A taxa de subida na altura do líquido quando a válvula de influxo está aberta é consideravelmente maior do que a taxa de descida na altura do líquido quando a válvula está fechada.

Solução. Suponha que o nível do líquido está baixando e que a válvula de influxo está fechada. Quando o nível atinge o nível AA' na Fig. 11.24(a), o limite inferior da abertura diferencial, os contatos da chave de mercúrio serão fechados e a válvula de influxo será aberta, permitindo a entrada de líquido no tanque. A altura do líquido começará a aumentar quando a válvula de influxo é aberta. No começo haverá uma taxa de subida alta. À medida que o nível do tanque sobe, o efluxo aumentará devido à maior altura de líquido. O resultado é um influxo total menor no tanque. Quando a altura alcança o nível BB' na Fig. 11.24(a), os contatos da chave de mercúrio serão abertos e a válvula de influxo será fechada. A altura do líquido então começará a cair.

A Fig. 11.25(a) mostra a curva da altura do líquido em função do tempo sob a Condição 1, quando a taxa de subida na altura com a válvula de influxo aberta é consideravelmente menor do que a taxa de descida na altura quando a válvula está fechada. Aqui o tempo de permanên-



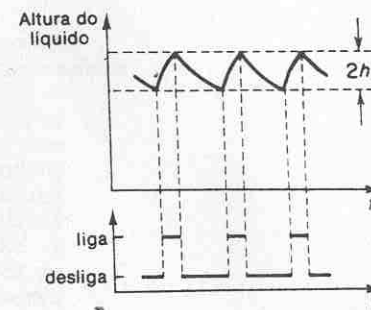


Fig. 11.25 (a) Curva da altura do líquido em função do tempo sob a Condição 1; (b) curva da altura do líquido em função do tempo sob a Condição 2.

cia no estado ligado é consideravelmente maior do que o tempo desligado. A Fig. 11.25(b) mostra uma curva de altura de líquido em função do tempo sob a Condição 2. O tempo ligado é consideravelmente menor do que o tempo desligado.

Em qualquer das condições, o nível oscila em torno do valor desejado. Portanto o sistema apresenta comportamento de ciclo limite. Em qualquer caso, o influxo médio é igual ao efluxo médio. As taxas de influxo-efluxo determinam a forma da curva do nível em função do tempo. Se a altura sobe com uma taxa que é igual à de descida com a válvula fechada, então o tempo ligado e o tempo desligado serão iguais.

Note que alargando a abertura da histerese *(gap)*, haverá operação menos freqüente da chave de mercúrio e da válvula operada por solenóide. Isto normalmente significa vida mais longa para o equipamento. A desvantagem de se operar com menor freqüência é que a variação da altura do líquido no tanque se torna maior.

Finalmente, deve-se notar que o controle liga-desliga oferece a melhor economia, a maior sensibilidade e facilidade de manutenção. Portanto, se um ciclo limite de amplitude pequena é permitido em uma aplicação particular, então o uso de qualquer outro tipo de controle seria um erro.

Problema A.11.2 Obtenha a função descritiva para a não linearidade liga-desliga com zona morta vista na Fig. 11.26.

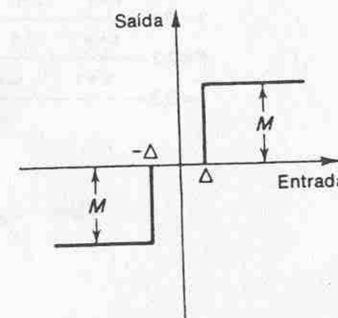


Fig. 11.26 Curva da característica de entrada-saída para a não linearidade tipo liga-desliga com zona morta.

Solução. A Fig. 11.27 mostra as formas de onda da entrada e da saída para um elemento com a dada não linearidade.

A saída do elemento não linear para $0 \leq \omega t \leq \pi$ é dada por

$$
\begin{aligned}
y(t) &= 0 \quad \text{para } 0 < t < t_1 \\
&= M \quad \text{para } t_1 < t < \frac{\pi}{\omega} - t_1 \\
&= 0 \quad \text{para } \frac{\pi}{\omega} - t_1 < t < \pi
\end{aligned}
$$

A forma de onda da saída é uma função ímpar. Portanto a componente harmônica fundamental da saída $y(t)$ é dada por

$$y_1(t) = Y_1 \operatorname{sen} \omega t$$

onde

$$
\begin{aligned}
Y_1 &= \frac{1}{\pi} \int_0^{2\pi} y(t) \operatorname{sen} \omega t \, d(\omega t) \\
&= \frac{4}{\pi} \int_0^{\pi/2} y(t) \operatorname{sen} \omega t \, d(\omega t) \\
&= \frac{4}{\pi} \int_{\omega t_1}^{\pi/2} M \operatorname{sen} \omega t \, d(\omega t) \\
&= \frac{4M}{\pi} \cos \omega t_1
\end{aligned}
$$

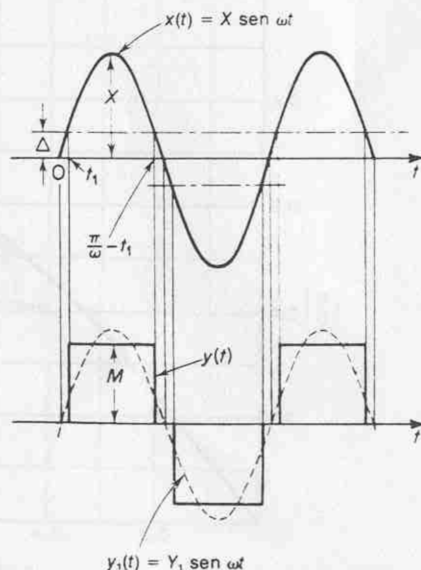


Fig. 11.27 Formas de onda de entrada e saída para a não linearidade tipo liga-desliga com zona morta.

Como $\operatorname{sen} \omega t_1 = \Delta/X$, obtemos

$$\cos \omega t_1 = \sqrt{1 - \left(\frac{\Delta}{X}\right)^2}$$

Portanto

$$y_1(t) = \frac{4M}{\pi} \sqrt{1 - \left(\frac{\Delta}{X}\right)^2} \operatorname{sen} \omega t$$

A função descritiva para um elemento liga-desliga com zona morta é

$$N = \frac{4M}{\pi X} \sqrt{1 - \left(\frac{\Delta}{X}\right)^2}$$

A função descritiva neste caso é uma função apenas da amplitude de entrada X. A Fig. 11.28 mostra um gráfico de $\Delta N/M$ em função de Δ/X.

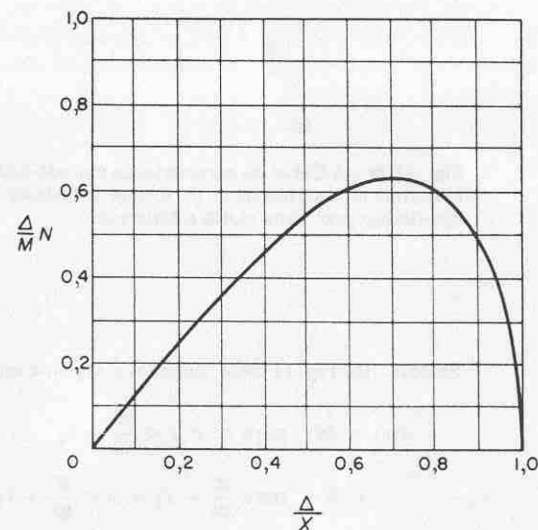


Fig. 11.28 Função descritiva para a não linearidade tipo liga-desliga com zona morta.

Problema A.11.3 A Fig. 11.29(a) mostra a curva de característica entrada-saída para uma não linearidade tipo liga-desliga com zona morta e histerese. A Fig. 11.29(b) mostra as formas de onda da entrada e saída de uma chave de contato com esta não linearidade.

A chave não se fecha até que a entrada exceda o valor $\Delta + h$. A chave permanece fechada até que a entrada se torna menor do que $\Delta - h$. Na região entre $\Delta - h$ e $\Delta + h$, a saída depende da história passada da entrada. A chave abre ou fecha de forma semelhante para uma entrada negativa.

Obtenha a função descritiva para a chave com esta não linearidade.

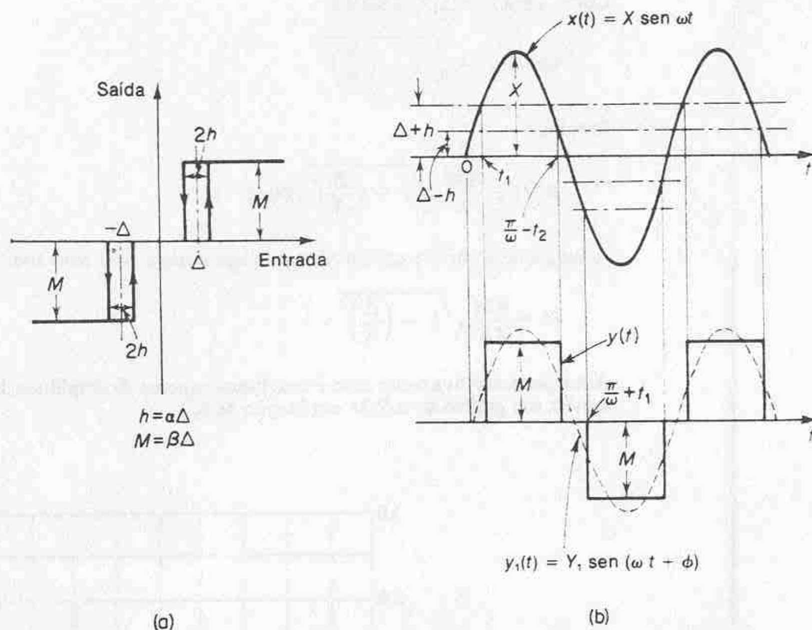


Fig. 11.29 (a) Curva da característica entrada-saída para a não linearidade tipo liga-desliga com zona morta e histerese; (b) formas de onda da entrada e saída para a não linearidade tipo liga-desliga com zona morta e histerese.

Solução. Da Fig. 11.29(b), obtemos a seguinte equação:

$$y(t) = M \quad \text{para } t_1 < t < \frac{\pi}{\omega} - t_2$$

$$= 0 \quad \text{para } \frac{\pi}{\omega} - t_2 < t < \frac{\pi}{\omega} + t_1$$

onde t_1 e t_2 são definidos por

$$\operatorname{sen} \omega t_1 = \frac{\Delta + h}{X}$$

$$\operatorname{sen} \omega t_2 = \frac{\Delta - h}{X}$$

Como visto na Fig. 11.29(b), a saída está atrasada em relação à entrada. A componente harmônica fundamental da saída pode ser obtida como segue:

$$y_1(t) = A_1 \cos \omega t + B_1 \operatorname{sen} \omega t$$

onde

$$A_1 = \frac{2}{\pi} \int_0^{\pi} y(t) \cos \omega t \, d(\omega t) = \frac{2}{\pi} \int_{\omega t_1}^{\pi - \omega t_2} M \cos \omega t \, d(\omega t) = -\frac{4hM}{\pi X}$$

$$B_1 = \frac{2}{\pi} \int_0^{\pi} y(t) \operatorname{sen} \omega t \, d(\omega t)$$

$$= \frac{2}{\pi} \int_{\omega t_1}^{\pi - \omega t_2} M \operatorname{sen} \omega t \, d(\omega t)$$

$$= \frac{2M}{\pi} \left[\sqrt{1 - \left(\frac{\Delta - h}{X}\right)^2} + \sqrt{1 - \left(\frac{\Delta + h}{X}\right)^2} \right]$$

Vamos definir

$$h = \alpha \, \Delta, \qquad M = \beta \, \Delta$$

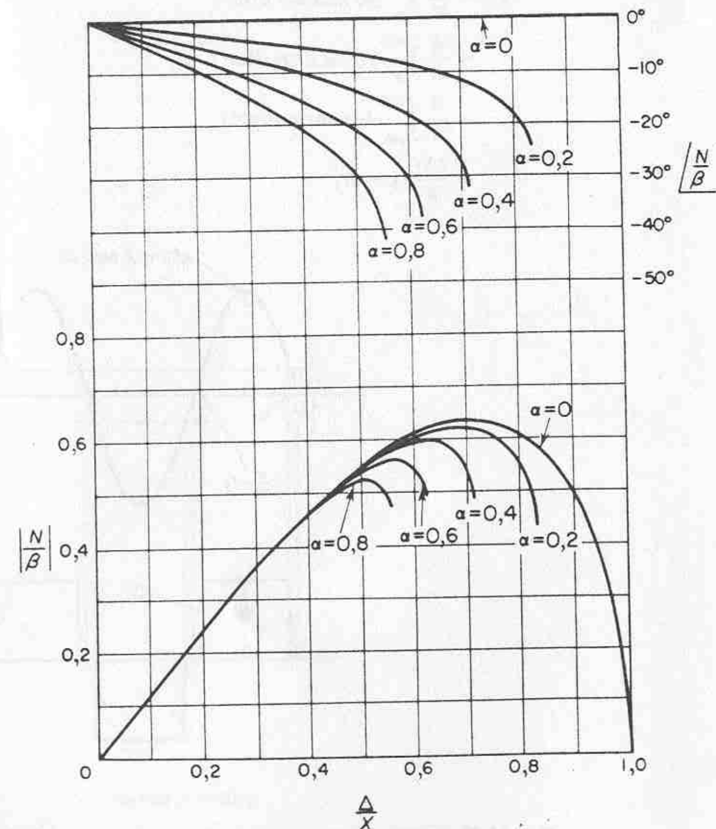


Fig. 11.30 Função descritiva para a não linearidade tipo liga-desliga com zona morta e histerese.

Os valores de α e β são constantes para uma dada não linearidade tipo liga-desliga com zona morta e histerese. Então,

$$\frac{A_1}{X} = -\frac{4\alpha\beta}{\pi}\left(\frac{\Delta}{X}\right)^2$$

$$\frac{B_1}{X} = \frac{2\beta}{\pi}\frac{\Delta}{X}\left[\sqrt{1-\left(\frac{\Delta}{X}\right)^2(1-\alpha)^2} + \sqrt{1-\left(\frac{\Delta}{X}\right)^2(1+\alpha)^2}\right]$$

Portanto a função descritiva para esta não linearidade é

$$N = \sqrt{\left(\frac{A_1}{X}\right)^2 + \left(\frac{B_1}{X}\right)^2} \angle \tan^{-1}\left(\frac{A_1}{B_1}\right)$$

Esta função descritiva é uma grandeza complexa. A Fig. 11.30 mostra gráficos de $|N/\beta|$ em função de Δ/X e $\angle N/\beta$ em função de Δ/X.

Problema A.11.4 O sistema visto na Fig. 11.31 apresenta um ciclo limite com a freqüência de oscilação de 5,9 rad/s, como visto na Fig. 11.32. Queremos diminuir a freqüência do ciclo limite para 4 rad/s. Determine a modificação necessária no ganho de $G(s)$, supondo que o elemento não linear é prefixado.

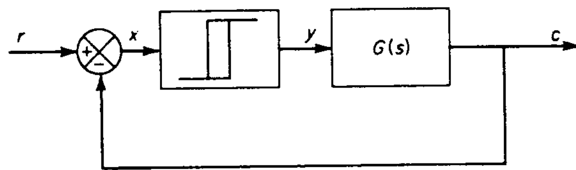


Fig. 11.31 Sistema de controle não linear.

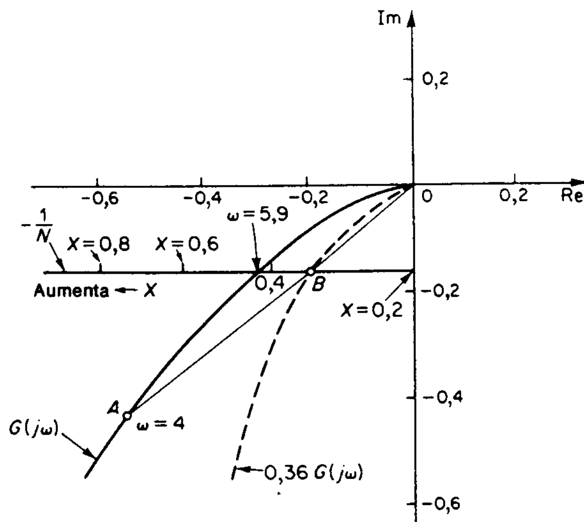
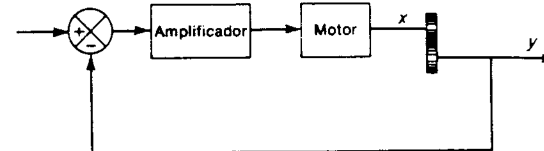
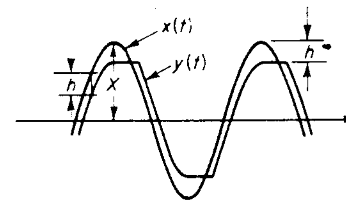


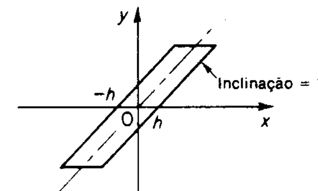
Fig. 11.32 Gráfico de $-1/N$ e $G(j\omega)$.



(a)



(b)



(c)

Fig. 11.33 (a) Diagrama em blocos de um servomecanismo com folga; (b) curvas características da não linearidade tipo folga; (c) curva da característica entrada-saída para uma não linearidade tipo folga (ou histerese).

Solução. Da Fig. 11.32, $\overline{OB}/\overline{OA}$ é determinado como sendo 0,36. Portanto, se o valor do ganho de $G(s)$ é diminuído para 36% do valor original, a freqüência do novo ciclo limite será 4 rad/s. A amplitude também diminui, de 0,42 para 0,35.

Problema A.11.5 A Fig. 11.33(a) mostra o diagrama em blocos de um servomecanismo consistindo em um amplificador, um motor e engrenagens. A posição da saída é realimentada para a entrada a fim de gerar o sinal de erro. Supõe-se que a inércia das engrenagens é desprezível comparada com a do motor, e também se supõe uma relação de engrenagens igual a um.

Em virtude da folga, os sinais $x(t)$ e $y(t)$ são relacionados como visto na Fig. 11.33(b). A Fig. 11.33(c) mostra a curva da característica de entrada-saída para a não linearidade tipo folga aqui considerada. A função descritiva para esta não linearidade é mostrada na Fig. 11.34.

Usando esta função descritiva, determine a amplitude e freqüência do ciclo limite quando a função de transferência da combinação amplificador-motor é dada por

$$\frac{5}{s(s+1)}$$

e a amplitude da folga é dada como unitária, ou $h = 1$.

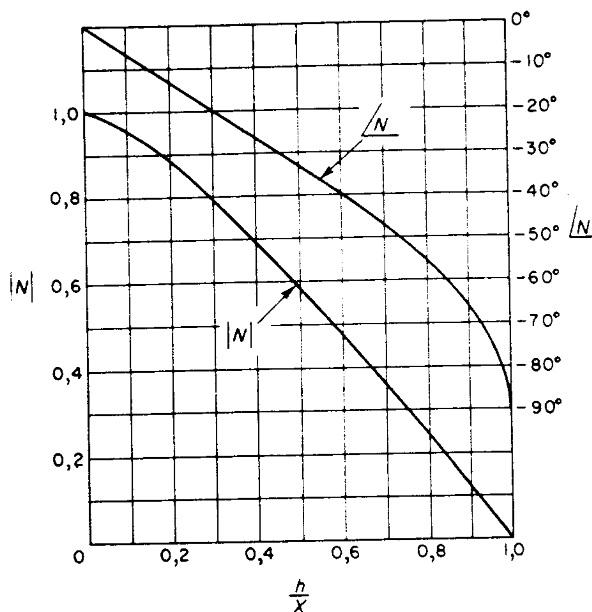


Fig. 11.34 Função descritiva para uma não linearidade tipo folga, ou uma não linearidade tipo histerese vista na Fig. 11.33(c).

Solução. Do enunciado do problema, o diagrama de blocos do sistema pode ser desenhado como visto na Fig. 11.35. A informação sobre a operação de ciclo limite do sistema pode ser facilmente obtida se fizermos um gráfico no domínio da freqüência.

A Fig. 11.36 mostra um gráfico dos lugares geométricos de $-1/N$ e de $G(j\omega)$ no diagrama do logaritmo do módulo em função da fase. Como visto no gráfico, há duas intersecções dos dois lugares geométricos. Aplicando o teste de estabilidade para o ciclo limite, como discutido na Seção 11.4, percebemos que o ponto A corresponde a um ciclo limite estável e que o ponto B corresponde a um ciclo limite instável. O ciclo limite estável tem uma freqüência de 1,6 rad/s e uma amplitude de 2. (O ciclo limite estável não pode ocorrer fisicamente.)

Para evitar comportamento de ciclo limite, o ganho do amplificador deve ser diminuído suficientemente para que o lugar geométrico de $G(j\omega)$ se localize bem abaixo do lugar geométrico de $-1/N$ no diagrama do log do módulo em função da fase. Se os dois lugares geométricos são tangentes, a precisão da análise por função descritiva é baixa, e podemos esperar um ciclo limite ou uma oscilação amortecida lentamente.

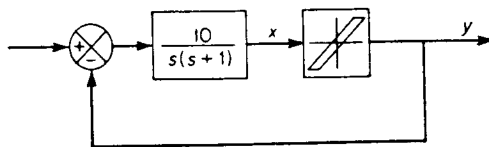


Fig. 11.35 Representação em diagrama em blocos para o servomecanismo visto na Fig. 11.33(a).

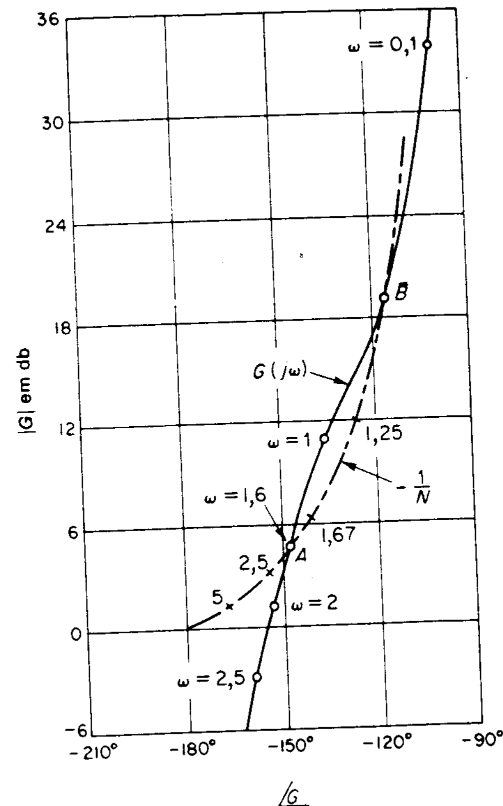


Fig. 11.36 Gráfico de $-1/N$ e $G(j\omega)$ do servomecanismo visto na Fig. 11.35.

PROBLEMAS

Problema B.11.1 Para o sistema visto na Fig. 11.37, determine a amplitude e freqüência do ciclo limite.

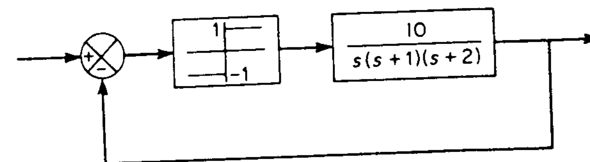


Fig. 11.37 Sistema de controle não linear.

Problema B.11.2 Determine a função descritiva para o elemento não linear descrito por

$$y = x^3$$

onde

x = entrada para o elemento não linear (sinal senoidal)
y = saída do elemento não linear

Problema B.11.3 Determine a estabilidade do sistema visto na Fig. 11.38.

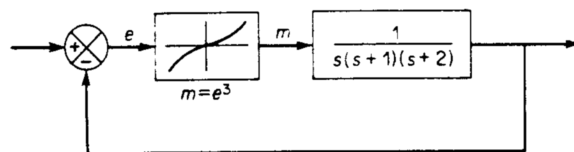


Fig. 11.38 Sistema de controle não linear.

Problema B.11.4 Determine a amplitude e a freqüência do ciclo limite do sistema visto na Fig. 11.39.

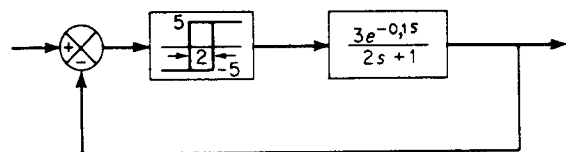


Fig. 11.39 Sistema de controle não linear.

Problema B.11.5 Determine a equação da função descritiva N para a não linearidade de histerese vista na Fig. 11.33(c).

Problema B.11.6 Referindo-se às formas de onda da entrada e saída para a não linearidade tipo saturação vista na Fig. 11.13(b), obtenha a amplitude da componente terceira harmônica da saída e faça um gráfico de Y_3/Y_1 como função de S/X, onde Y_1 é a amplitude da componente harmônica fundamental e Y_3 é a amplitude da componente terceira harmônica.

Problema B.11.7* Considere um elemento não linear cuja característica de entrada-saída é definida por

$$y = b_1x + b_3x^3 + b_5x^5 + b_7x^7 + \cdots$$

onde
x = entrada para o elemento não linear (sinal senoidal)
y = saída do elemento não linear

Mostre que a função descritiva para esta não linearidade pode ser dada por

$$N = b_1 + \tfrac{3}{4}b_3X^2 + \tfrac{5}{8}b_5X^4 + \tfrac{35}{64}b_7X^6 + \cdots$$

onde X é a amplitude da senóide de entrada $x = X \text{ sen } \omega t$.

*Referência W-1.

12

Análise por Plano de Fase

12.1 INTRODUÇÃO

Considere um sistema de segunda-ordem descrito pela seguinte equação diferencial ordinária:

$$\ddot{x} + f(x, \dot{x}) = 0$$

onde $f(x, \dot{x})$ é uma função linear ou não linear de x e $\dot{x}$. A solução temporal deste sistema pode ser ilustrada através de um gráfico de $x(t)$ em função de t. Ela também pode ser ilustrada fazendo-se o gráfico de $\dot{x}(t)$ em função de $x(t)$ usando t como parâmetro.

Se tomamos x e $\dot{x}$ como as coordenadas de um plano, a cada estado do sistema corresponde um ponto neste plano. Quando t varia, este ponto descreve uma curva no plano $x - \dot{x}$, indicando a história do sistema. Tal curva é chamada uma *trajetória*.

A representação geométrica do comportamento do sistema em termos de trajetórias é chamada uma representação de plano de fase da dinâmica do sistema. Embora o diagrama do plano de fase forneça uma visualização clara das trajetórias para sistemas de segunda-ordem, normalmente é difícil visualizar ou construir trajetórias para sistemas de terceira-ordem. Para sistemas de ordem maior que três, é impossível visualizar as trajetórias; entretanto, noções sobre o movimento de um ponto representativo em espaço de duas dimensões podem ser conceitualmente estendidas para espaços n-dimensionais.

Este capítulo apresenta o método do plano de fase para análise de sistemas de segunda-ordem. O método é especialmente útil no tratamento de sistemas com fortes não linearidades. Ao contrário do método da função descritiva, o método do plano de fase não é restrito a pequenas não linearidades. A familiaridade com este método pode nos permitir resolver problemas de controle não linear específicos, bem como entender as características fundamentais de certos sistemas não lineares. Embora os sistemas que podem ser analisados por este método sejam praticamente limitados a sistemas de primeira e segunda-ordem, os resultados obtidos estudando-se os efeitos de uma não linearidade na resposta de um sistema de segunda-ordem fornecem considerável intuição no comportamento de sistemas de ordem superior tendo a mesma não linearidade. Neste sentido, o método do plano de fase é bastante útil para analisar sistemas de controle nao lineares. Também é útil para sintetizar tais sistemas. Em particular, o controle ótimo de sistemas de segun-

da-ordem pode ser estudado convenientemente no plano de fase (veja Cap. 16). Note, entretanto, que as limitações na ordem do sistema diminuem a utilidade deste método no estudo de sistemas de controle ótimo de sistemas de ordem superior.

Método do plano de fase. O método do plano de fase introduzido por Poincare é um método de se obter graficamente a solução das seguintes equações diferenciais simultâneas de primeira-ordem:

$$\frac{dx_1}{dt} = f_1(x_1, x_2) \tag{12.1}$$

$$\frac{dx_2}{dt} = f_2(x_1, x_2) \tag{12.2}$$

onde $f_1(x_1, x_2)$ e $f_2(x_1, x_2)$ são funções lineares ou não lineares das variáveis x_1 e x_2, respectivamente. As Eqs. (12.1) e (12.2) são chamadas autônomas, o que significa que a variável independente t aparece apenas na forma de derivadas. (Portanto, em um sistema autônomo, nem as forças nem as condições de contorno variam com o tempo.)

O plano com coordenadas retangulares x_1 e x_2 é chamado plano de fase, ou plano de estados. (O plano de fase ou plano de estados é um espaço de estados bidimensional. Discutiremos em detalhes o espaço de estados no Cap. 14.)

Freqüentemente as Eqs. (12.1) e (12.2) assumem a seguinte forma simplificada:

$$\frac{dx_1}{dt} = x_2$$

$$\frac{dx_2}{dt} = f(x_1, x_2)$$

Se definirmos $x_1 = x$, então $x_2 = \dot{x}$. O plano de fase mais comum é o plano $x - \dot{x}$. Neste capítulo, a não ser que mencionemos o contrário, supomos que o plano de fase é o plano $x - \dot{x}$.

A análise por plano de fase dos sistemas de Eqs. (12.1) e (12.2) fornece uma visão geral das soluções para quaisquer condições iniciais possíveis. No campo de sistemas de controle, o método do plano de fase é particularmente adequado na análise e síntese de sistemas de segunda-ordem sujeitos a condições iniciais e/ou entradas aperiódicas, como entradas em degrau, entradas em rampa, entradas em pulso e entradas impulsivas.

Trajetórias no plano de fase. Do teorema fundamental da unicidade da solução de equações diferenciais simultâneas, sabemos que a solução das Eqs. (12.1) e (12.2) com uma dada condição inicial é única, desde que $f_1(x_1, x_2)$ e $f_2(x_1, x_2)$ nas Eqs. (12.1) e (12.2) são analíticas. (Uma função é analítica em um certo ponto se é possível obter uma expansão em série de Taylor da função ao redor do dado ponto.) Este resultado de unicidade não se aplica aos pontos onde $f_1(x_1, x_2) = 0$ e $f_2(x_1, x_2) = 0$, simultaneamente. Tais pontos são chamados pontos singulares. Pontos singulares são pontos de equilíbrio. Qualquer outro ponto no plano de fase é chamado um ponto ordinário.

Se não há outros pontos de equilíbrio nas vizinhanças de um dado ponto de equilíbrio, então este ponto de equilíbrio é chamado de um ponto isolado. Embora muitos sistemas reais envolvam apenas pontos de equilíbrio isolados, há alguns

casos em que isto não é verdade. Por exemplo, para o sistema

$$\ddot{x} + \dot{x} = 0$$

todos os pontos no eixo x são pontos de equilíbrio, e, portanto, os pontos de equilíbrio não são isolados.

Obtenção de equações diferenciais de primeira-ordem a partir de sistemas de segunda-ordem. A eliminação da variável independente t das Eqs. (12.1) e (12.2) fornece

$$\frac{dx_2}{dx_1} = \frac{f_2(x_1, x_2)}{f_1(x_1, x_2)} \tag{12.3}$$

A Eq. (12.3) é uma equação diferencial de primeira-ordem relacionando x_1 e x_2, e, de fato, esta equação dá a inclinação da tangente à trajetória passando pelo ponto (x_1, x_2).

O estado do sistema dado pelas Eqs. (12.1) e (12.2) [ou Eq. (12.3)] pode ser determinado a qualquer instante t pelos valores de x_1 e x_2.

A solução da Eq. (12.3) pode ser escrita

$$x_2 = \phi(x_1) \tag{12.4}$$

A Eq. (12.4) representa uma curva no plano de fase e indica um movimento de um ponto representativo sobre a curva. A curva-solução, ou trajetória, que é um gráfico de x_2 em função de x_1, é uma curva integral do sistema representado pela Eq. (12.3). A trajetória não mostra informação temporal de forma explícita. Se necessário, a trajetória pode ser graduada em unidades de tempo.

Gráficos de plano de fase. Uma família de trajetórias é chamada um gráfico de plano de fase (*phase-plane portrait*). A condição inicial determina a localização inicial de um ponto representativo na trajetória. Quando o tempo aumenta, o ponto representativo se move ao longo da trajetória. A representação por plano de fase de um sistema autônomo mostra a totalidade de todos os possíveis estados do sistema, e portanto a natureza da resposta do sistema é mostrada diretamente no gráfico de plano de fase. Como há uma e apenas uma trajetória passando por qualquer dado ponto ordinário no plano de fase, as trajetórias geradas por todas as possíveis condições iniciais não se cruzam, exceto em pontos singulares. Em pontos singulares, dx_2/dx_1 é indeterminado pois é da forma zero sobre zero. Uma infinidade de trajetórias podem-se aproximar ou se afastar de um ponto singular.

Para um sistema tendo uma não linearidade de dois valores, tal como a não linearidade tipo histerese, o sistema perde a sua analiticidade. Em tal caso, entretanto, pode ser possível dividir a região em sub-regiões em que o sistema é analítico, e portanto o método do plano de fase pode ser aplicado. Então a solução completa pode ser obtida conectando-se soluções analíticas por trechos. Deve-se notar que quando a resposta do sistema é definida por duas ou mais equações diferenciais de segunda-ordem, as trajetórias podem se cruzar.

Exemplo 12.1 Considere o sistema de primeira-ordem descrito por

$$\dot{x} = -x \tag{12.5}$$

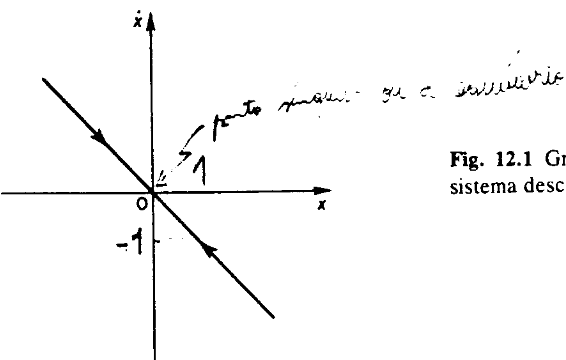


Fig. 12.1 Gráfico de plano de fase de um sistema descrito por $\dot{x} = -x$.

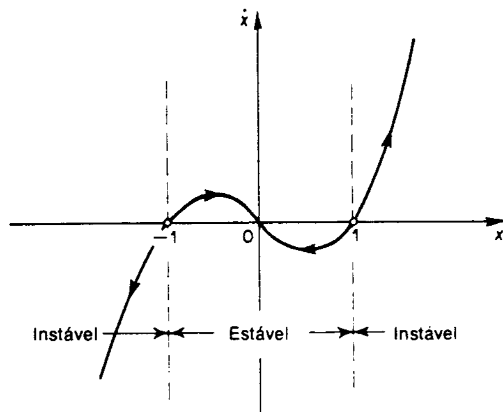


Fig. 12.2 Gráfico de plano de fase de um sistema descrito por $\dot{x} = -x + x^3$.

No plano de fase, ou plano x–$\dot{x}$, a Eq. (12.5) representa uma linha reta, como visto na Fig. 12.1. Para qualquer condição inicial $x(0)$, o sistema volta para seu ponto singular, a origem, após um tempo infinito. A condição inicial $x(0)$ determina o ponto de partida da trajetória.

Considere em seguida o seguinte sistema de primeira-ordem:

$$\dot{x} = -x + x^3$$

A trajetória é vista na Fig. 12.2. Pode-se ver que a trajetória é dividida em três partes, uma parte estável e duas partes instáveis. Isto é, se $x(0) > 1$, então $x(\infty) \to \infty$. Se $1 > x(0) > -1$, então $x(\infty) \to 0$. Se $x(0) < -1$, então $x(\infty) \to -\infty$.

Exemplo 12.2 Considere o sistema representado pela seguinte equação:

$$\ddot{x} + \dot{x} + x = 0 \qquad (12.6)$$

A Fig. 12.3 mostra um gráfico de plano de fase para este sistema. A Eq. (12.6) resulta em

$$\frac{\ddot{x}}{\dot{x}} = \frac{\dfrac{d\dot{x}}{dt}}{\dfrac{dx}{dt}} = \frac{d\dot{x}}{dx} = \frac{-\dot{x} - x}{\dot{x}}$$

Na origem

$$\frac{d\dot{x}}{dx} = \frac{0}{0}$$

A origem é portanto um ponto singular, isto é, o sistema está em equilíbrio na origem. Quando as condições iniciais da Eq. (12.6) são $x(0) = 0$ e $\dot{x}(0) = 10$, a trajetória correspondente é a curva $ABCDO$ na Fig. 12.3. De forma semelhante, para $x(0) = 2$ e $\dot{x}(0) = -7$, a trajetória é a curva EFO.

Se um ponto representativo está na metade superior do plano x–$\dot{x}$ (por exemplo, o ponto B na Fig. 12.3), o ponto se move para a direita em uma trajetória quando o tempo aumenta, pois uma velocidade positiva ($\dot{x} > 0$) corresponde a um aumento no valor de x com o tempo. De forma similar, se um ponto representativo está na metade inferior do plano x–$\dot{x}$ (por exemplo, o ponto D na Fig. 12.3), então o ponto se move para a esquerda em uma trajetória quando o tempo aumenta, uma vez que uma velocidade negativa ($\dot{x} < 0$) corresponde a um decréscimo em x com o tempo. Portanto, movimentos ao longo de uma trajetória no plano x–$\dot{x}$ se fazem na direção horária. Quando a trajetória cruza o eixo x, a velocidade $\dot{x}$ é nula. Portanto, a trajetória cruzará o eixo x perpendicularmente.

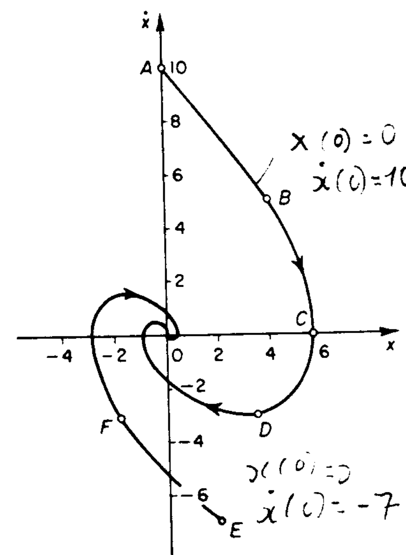


Fig. 12.3 Gráfico de plano de fase de um sistema descrito por $\ddot{x} + \dot{x} + x = 0$.

12.2 MÉTODOS PARA CONSTRUIR TRAJETÓRIAS

Na análise de plano de fase de sistemas de segunda-ordem as trajetórias podem ser construídas analiticamente, graficamente ou experimentalmente. Nesta seção apresentaremos métodos analíticos e gráficos para construir trajetórias.

Os métodos analíticos são úteis para sistemas cujas equações diferenciais são simples e lineares por trechos. Obviamente, se a análise do sistema é de interesse primordial e se a solução temporal da equação diferencial do sistema pode ser

obtida facilmente por métodos analíticos, há pouca justificativa de se fazer um gráfico de plano de fase do sistema. Entretanto, métodos analíticos para se obter as equações das trajetórias se mostram úteis em casos onde as expressões analíticas das trajetórias de equações diferenciais lineares simples são necessárias para sintetizar sistemas não lineares ou para auxiliar em uma demonstração de existência de uma curva fechada no plano de fase etc.

Os métodos gráficos são úteis quando é cansativo, difícil ou impossível de se resolver a dada equação diferencial analiticamente. Métodos gráficos podem ser aplicados tanto para equações lineares quanto não lineares. Estes métodos fornecem diretamente gráficos de trajetórias no plano de fase. Como a maioria das equações diferenciais não lineares de segunda-ordem não podem ser resolvidas analiticamente, então, é para tais equações que os métodos gráficos têm muito que oferecer.

Todos os métodos gráficos são essencialmente baseados em procedimentos tipo passo-a-passo. O número de passos necessários para se obter uma trajetória depende do tamanho dos incrementos utilizados. Estes incrementos devem ser de magnitude adequada de tal forma que ocorram apenas variações relativamente pequenas durante cada incremento. Eles não devem ser nem muito pequenos nem muito grandes. Se os incrementos forem muito pequenos, os cálculos podem levar muito tempo.

A precisão de métodos gráficos depende da maneira com que se faz a construção. Obviamente, o grau de precisão pode ser aumentado até um certo ponto, à medida que o tamanho da figura é aumentado. Deve ficar claro que se os incrementos são muito grandes, a precisão provavelmente será ruim. No entanto devemos notar que, mesmo que os incrementos sejam pequenos, se o número de incrementos for muito grande, a precisão provavelmente também será ruim. Isto ocorre por causa de erros inevitáveis, pois, mesmo que eles sejam pequenos a cada passo, tendem a se acumular e podem causar uma falta de precisão na porção final da solução.

Os dois métodos gráficos discutidos neste capítulo são o método das isóclinas e o método delta. Quaisquer parâmetros na dada equação devem, obviamente, ter associados valores numéricos antes que os métodos gráficos possam ser aplicados.

Também existem técnicas experimentais para a obtenção das trajetórias. Uma vez que montamos as equações para a análise do plano de fase, podemos em seguida usar um computador analógico e um osciloscópio ou traçador xy. Por exemplo, as trajetórias podem ser visualizadas na tela aplicando-se às placas defletoras de um osciloscópio tensões proporcionais a x para a deflexão horizontal e a $\dot{x}$ para a deflexão vertical. Tal técnica experimental para a obtenção das trajetórias permite fazer uma análise de plano de fase geral para o problema além de economizar tempo que seria necessário no cálculo manual para construir as trajetórias.

Simetria nos gráficos de plano de fase. Em certos casos, o gráfico de plano de fase pode ser simétrico em torno do eixo x, do eixo $\dot{x}$, ou de ambos. A simetria no gráfico de plano de fase pode ser facilmente determinada a partir das equações diferenciais originais ou da equação representando a trajetória.

Simetria em torno do eixo x: Considere a seguinte equação:

$$\ddot{x} + f(x, \dot{x}) = 0$$

ou

$$\frac{d\dot{x}}{dx} = -\frac{f(x, \dot{x})}{\dot{x}}$$

Para que as trajetórias sejam simétricas em torno do eixo x, a inclinação $d\dot{x}/dx$ deve ser igual mas de sinal oposto para $\dot{x} > 0$ e $\dot{x} < 0$ para todo x. Desta condição pode-se mostrar que, se o gráfico de plano de fase é simétrico em torno do eixo x, então

$$f(x, \dot{x}) = f(x, -\dot{x})$$

ou seja, $f(x, \dot{x})$ deve ser uma função par de $\dot{x}$. Portanto, o gráfico de plano de fase de uma equação diferencial que não contém $\dot{x}$ é sempre simétrico em torno do eixo x.

Se o gráfico de plano de fase é simétrico em torno do eixo x, a única diferença entre o gráfico no semiplano superior daquele no semiplano inferior é a direção do movimento. Isto é, o movimento de um ponto representativo no semiplano superior é para a direita, e no semiplano inferior é para a esquerda. Portanto, para este caso, pode não ser necessário levantar o gráfico de plano de fase para o semiplano inferior.

Simetria em torno do eixo $\dot{x}$: Para simetria em torno do eixo $\dot{x}$, a inclinação $d\dot{x}/dx$ deve ser igual mas de sinal oposto para $x > 0$ e $x < 0$ para qualquer $\dot{x}$. Isto requer que

$$f(x, \dot{x}) = -f(-x, \dot{x})$$

ou $f(x, \dot{x})$ deve ser uma função ímpar de x.

Simetria em torno tanto do eixo x como do eixo $\dot{x}$. Para o gráfico do plano de fase ser simétrico em torno de ambos os eixos x e $\dot{x}$, as condições para os dois casos anteriores devem ser satisfeitas simultaneamente. Isto requer que

$$f(-x, \dot{x}) = -f(x, -\dot{x})$$

Métodos analíticos para construir trajetórias. Há dois métodos para a determinação analítica das equações das trajetórias. Um método é integrar a Eq. (12.3) para obter a seguinte equação da trajetória:

$$x_2 = \phi(x_1) \tag{12.7}$$

Este método só é aplicável se a Eq. (12.3) puder ser integrada. Uma vez que a Eq. (12.7) é obtida, a relação entre x_1 e x_2 pode ser diretamente posta num gráfico. O outro método é obter x_1 e x_2 como funções de t e então eliminar t dessas duas equações. Este método pode ser aplicado se a eliminação de t das Eqs. (12.1) e (12.2) pode ser feita sem muita dificuldade.

Exemplo 12.3 Considere a seguinte equação diferencial de segunda-ordem simples:

$$\ddot{x} + \omega^2 x = 0$$

Usando a relação $\ddot{x} = d\dot{x}/dt = \dot{x}\,d\dot{x}/dx$, podemos mudar a equação para a seguinte forma:

$$\dot{x}\frac{d\dot{x}}{dx} + \omega^2 x = 0 \tag{12.8}$$

O primeiro método analítico é integrar a Eq. (12.8). Isto pode ser feito facilmente. e a equação para a trajetória é obtida como

$$\frac{\dot{x}^2}{\omega^2} + x^2 = A^2 \qquad (12.9)$$

onde A é uma constante, determinada a partir das condições iniciais.

Para um sistema simples como este, o segundo método analítico também pode ser aplicado. A solução $x(t)$ pode ser determinada como

$$x(t) = A\cos(\omega t + \alpha) \qquad (12.10)$$

onde A e α são constantes a serem determinadas das condições iniciais. Derivando a Eq. (12.10) com respeito a t, obtemos

$$\dot{x}(t) = -A\omega\,\text{sen}(\omega t + \alpha) \qquad (12.11)$$

Eliminando t das Eqs. (12.10) e (12.11), podemos obter a Eq. (12.9).

Exemplo 12.4 Considere a seguinte equação:

$$\ddot{x} = -M \qquad (M = \text{constante}) \qquad (12.12)$$

com as condições iniciais

$$x(0) = x_0, \qquad \dot{x}(0) = 0$$

A Eq. (12.12) pode ser escrita como

$$\dot{x}\frac{d\dot{x}}{dx} = -M \qquad (12.13)$$

Separando as variáveis e integrando a Eq. (12.13), obtém-se

$$\dot{x}^2 = 2(x_0 - x)M \qquad (12.14)$$

onde a constante de integração foi determinada das condições iniciais dadas. A Eq. (12.14) representa uma trajetória passando pelo ponto $(x_0, 0)$.

A Eq. (12.14) pode também ser obtida achando-se $x = \phi_1(t)$, $\dot{x} = \phi_2(t)$ e eliminando t destas duas equações. Usando as condições iniciais dadas, $\dot{x}(t)$ e $x(t)$ são determinadas como

$$\dot{x}(t) = -Mt \qquad (12.15)$$

$$x(t) = -\tfrac{1}{2}Mt^2 + x_0 \qquad (12.16)$$

Eliminando t das Eqs. (12.15) e (12.16), resulta na Eq. (12.14). Os gráficos de plano de fase do sistema correspondendo a $M = 1$ e $M = -1$ são vistos na Fig. 12.4. A trajetória particular correspondendo a $M = 1$ e passando pelo ponto $(x_0, 0)$ é mostrada como uma curva mais grossa.

Método gráfico de construção de trajetórias — o método das isóclinas. O método das isóclinas é usado para obter o campo de trajetórias graficamente, sem ter que

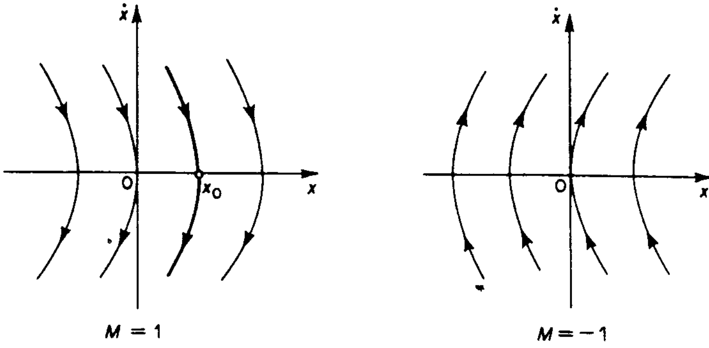


Fig. 12.4 Gráficos de plano de fase de um sistema descrito por $\ddot{x} = -M$ para $M = 1$ e $M = -1$.

resolver uma dada equação diferencial. O método das isóclinas é aplicável à seguinte equação diferencial de primeira-ordem:

$$\frac{dx_2}{dx_1} = \frac{f_2(x_1, x_2)}{f_1(x_1, x_2)} \qquad (12.17)$$

onde $f_1(x_1, x_2)$ e $f_2(x_1, x_2)$ são analíticas. Na Eq. (12.17), x_1 é considerada a variável independente e x_2 a variável dependente.

O lugar geométrico de uma inclinação constante da trajetória, isto é, o lugar geométrico de

$$\frac{dx_2}{dx_1} = \alpha = \text{constante} \qquad (12.18)$$

no plano de fase pode ser obtido das Eqs. (12.17) e (12.18). Ou seja, a seguinte equação:

$$f_2(x_1, x_2) = \alpha f_1(x_1, x_2)$$

fornece o lugar geométrico de inclinação constante α. O lugar geométrico de pontos onde as trajetórias têm uma dada inclinação é chamado de uma isóclina. Se desenharmos isóclinas correspondendo a vários valores de α, o campo de direções de tangentes às trajetórias pode ser obtido.

Quando construímos um gráfico de plano de fase pelo método das isóclinas, todo o plano de fase pode ser preenchido por pequenos segmentos de reta que fixam direções do campo, como visto na Fig. 12.5. Em muitos casos, as isóclinas para inclinação zero e infinito podem ser facilmente construídas. Como a inclinação é indeterminada em um ponto singular, não se pode desenhar nenhum segmento de reta neste ponto. A trajetória passando por um dado ponto ordinário qualquer no plano de fase pode ser construída desenhando-se uma curva contínua seguindo as direções do campo.

Como um exemplo, considere a equação diferencial

$$\ddot{x} + 2\zeta\omega\dot{x} + \omega^2 x = 0 \qquad (12.19)$$

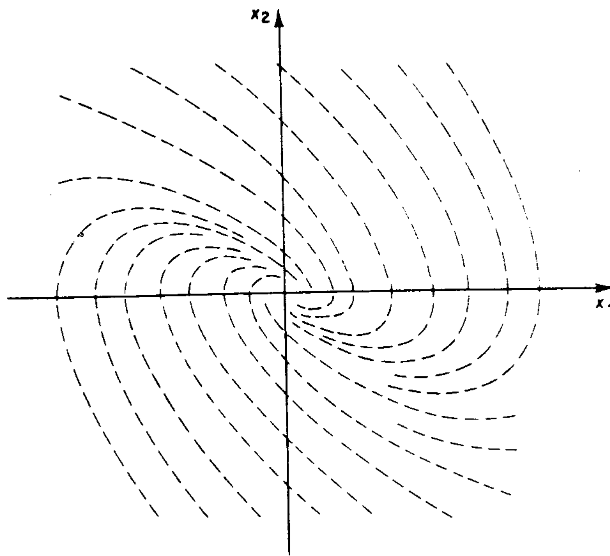


Fig. 12.5 Um exemplo de um diagrama de plano de fase preenchido com segmentos de reta curtos dando as direções do campo.

A Eq. (12.19) pode ser reescrita como segue:

$$\dot{x}\frac{d\dot{x}}{dx} - 2\zeta\omega\dot{x} - \omega^2 x = 0 \qquad (12.20)$$

Fazendo $d\dot{x}/dx = \alpha$, podemos escrever a Eq. (12.20) como

$$\frac{\dot{x}}{x} = \frac{-\omega^2}{2\zeta\omega + \alpha} \qquad (12.21)$$

A Eq. (12.21) representa uma isóclina, o lugar geométrico de tangente constante α, e é uma equação de uma reta. As isóclinas para equações lineares de segunda-ordem são linhas retas e passam pela origem do plano de fase.

A Fig. 12.6 mostra isóclinas e uma particular trajetória para o sistema dado pela Eq. (12.19) quando $\zeta = 0{,}5$ e $\omega = 1$. A trajetória começando no ponto A na Fig. 12.6 pode ser construída como segue: Quando a trajetória está na região delimitada pelas isóclinas correspondendo a $\alpha = -1$ e $\alpha = -1{,}2$, a inclinação da trajetória é aproximadamente $(-1 - 1{,}2)/2 = -1{,}1$. Portanto, se a reta com inclinação $-1{,}1$ é desenhada do ponto A, interceptando a isóclina correspondendo a $\alpha = -1{,}2$ no ponto B, a reta AB é aproximadamente uma parte da trajetória. Similarmente, a trajetória passando pelo ponto B tem uma inclinação de aproximadamente $(-1{,}2 - 1{,}4)/2 = -1{,}3$ na região delimitada pelas isóclinas correspondendo a $\alpha = -1{,}2$ e $\alpha = -1{,}4$. Então a reta BC com inclinação $-1{,}3$, interceptando a isóclina $\alpha = -1{,}4$ no ponto

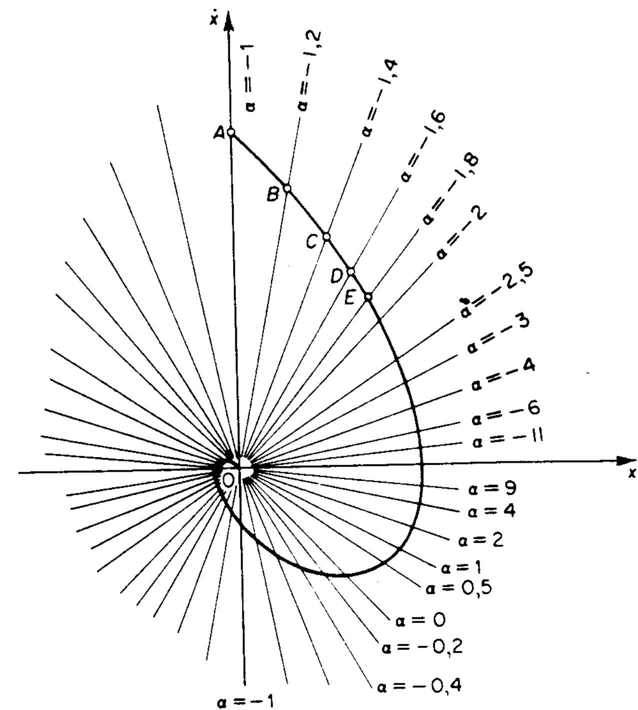


Fig. 12.6 Diagrama de plano de fase mostrando isóclinas e uma trajetória típica para um sistema descrito por $\ddot{x} + \dot{x} + x = 0$.

C, é uma parte da trajetória. A inclinação da trajetória entre os pontos C e D é $-1{,}5$. Construindo a trajetória desta forma, podemos obter a curva $ABCDEO$.

Para um sistema linear, há apenas uma trajetória passando por qualquer ponto ordinário no plano de fase. Portanto, a trajetória partindo do ponto D na Fig. 12.6 a curva DEO, uma parte da curva $ABCDEO$. Obviamente, a trajetória pode ser construída tanto na direção horária como anti-horária, a partir de qualquer ponto no plano de fase.

Deve-se notar que, ao se construir as trajetórias desta forma, podemos acumular erros que podem ocorrer durante o curso da construção, e a trajetória resultante pode não ser precisa. Para boa precisão, é conveniente primeiro preencher uma certa região do plano de fase por pequenos segmentos de reta que indicam direções do campo. Então constrói-se uma trajetória fazendo um esboço cuidadoso, seguindo as direções do campo. Deste forma, pode-se evitar a eventualidade de se acumular certos erros.

Lembre-se de que a precisão do resultado pode não ser boa se, como no caso de certas equações diferenciais não lineares, as inclinações das trajetórias variam rapidamente em certas regiões do plano de fase. A precisão deste método, em geral, depende do número de isóclinas utilizadas. Para uma razoável precisão, as isóclinas podem ser desenhadas a cada 5° ou 10°.

Para o sistema dado pela Eq. (12.19), a equação da isóclina, Eq. (12.21), pode ser escrita como

$$\alpha = -\left(2\zeta\omega + \frac{\omega^2}{\tan\theta}\right) \tag{12.22}$$

onde usamos a relação

$$\frac{\dot{x}}{x} = \tan\theta$$

e o ângulo θ é aquele visto na Fig. 12.7. Usando a Eq. (12.22), podemos desenhar isóclinas para qualquer valor de θ.

O método das isóclinas discutido aqui é conveniente quando as isóclinas são linhas retas. Se não o forem, o método delta, que será discutido a seguir, pode ser menos monótono para se construir a trajetória passando por um dado ponto.

Deve-se notar que a inclinação $\alpha = d\dot{x}/dx$ é uma quantidade com uma dimensão e seu valor depende da unidade de tempo. É muito conveniente usar as mesmas escalas ao longo dos eixos de x e $\dot{x}$ para que o valor de α e a correspondente inclinação geométrica da trajetória sejam os mesmos.

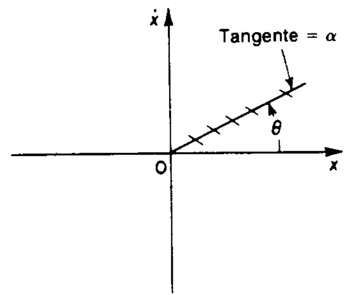


Fig. 12.7 Diagrama mostrando o ângulo θ da Eq. (12.22).

Exemplo 12.5 Obtenha um gráfico de plano de fase da seguinte equação usando o método das isóclinas:

$$\ddot{x} + a|\dot{x}| + x = 0 \qquad (a > 0) \tag{12.23}$$

Em geral, quando a equação do sistema envolve valores absolutos de x ou $\dot{x}$ ou ambos, a substituição da equação diferencial não linear por várias equações diferenciais lineares simplifica a construção das isóclinas.

Vamos substituir a Eq. (12.23) por

$$\begin{aligned} \ddot{x} + a\dot{x} + x &= 0 \quad \text{para } \dot{x} > 0 \\ \ddot{x} - a\dot{x} + x &= 0 \quad \text{para } \dot{x} < 0 \end{aligned} \tag{12.24}$$

No semiplano superior do plano x–$\dot{x}$, a equação das isóclinas é

$$\frac{\dot{x}}{x} = \frac{-1}{a + \alpha}$$

e no semiplano inferior é

$$\frac{\dot{x}}{x} = \frac{-1}{-a + \alpha}$$

A Fig. 12.8 mostra um gráfico de plano de fase, juntamente com isóclinas para o sistema da Eq. (12.23) quando $a = 1$. As linhas retas são isóclinas e os segmentos de reta curtos em cada linha indicam o campo de direções de tangentes às trajetórias. (Segmentos de reta curtos são colocados para que as trajetórias possam ser facilmente visualizadas.) Para este sistema, as trajetórias são curvas ovais concêntricas, e o movimento é periódico para quaisquer condições iniciais dadas, exceto na origem. Isto pode estar claro também pela Eq. (12.24) uma vez que a energia dissipada durante o meio ciclo (correspondendo a um período para o qual $\dot{x} > 0$) é a mesma energia fornecida ao sistema durante o outro meio-ciclo (correspondendo a um período para o qual $\dot{x} < 0$), e portanto a dissipação total de energia é nula para um ciclo. Neste exemplo, a natureza simétrica do gráfico de plano de fase pode ser prevista diretamente da Eq. (12.23). Como a substituição de $\dot{x}$ por $-\dot{x}$ não modifica a equação, ela é simétrica em torno do eixo x.

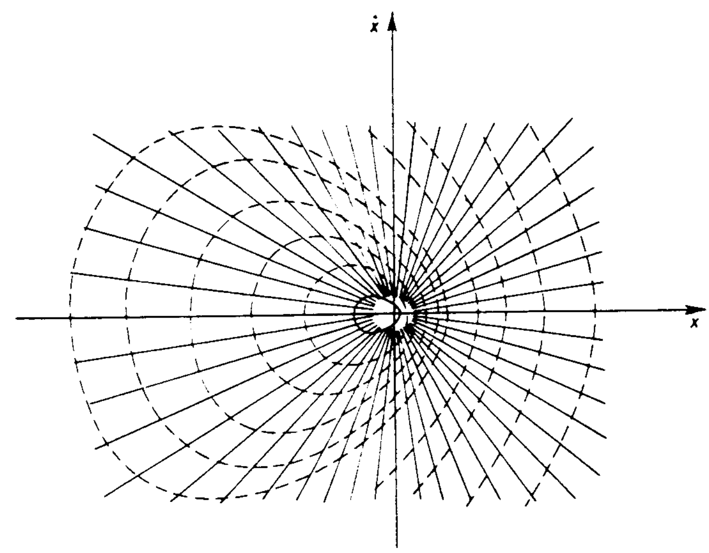


Fig. 12.8 Gráfico de plano de fase para um sistema descrito por $\ddot{x} + |\dot{x}| + x = 0$.

Método gráfico para construir trajetórias — o método delta. No método delta, a trajetória é obtida como uma seqüência de arcos circulares cujos centros se deslocam ao longo do eixo x. O método delta pode ser aplicado a equações da forma

$$\ddot{x} = -f(\dot{x}, x, t) \tag{12.25}$$

onde $f(\dot{x}, x, t)$ pode ser tanto linear como não linear e pode ser variante no tempo, mas deve ser contínua e unívoca. Ao aplicar este método, modificamos a Eq.

(12.25) para a seguinte forma:

$$\ddot{x} + \omega^2 x = -f(\dot{x}, x, t) + \omega^2 x \tag{12.26}$$

Um termo $\omega^2 x$ é adicionado a ambos os lados da Eq. (12.25). Este termo deve ser escolhido adequadamente de tal forma que os valores da função δ definida abaixo não sejam nem muito pequenos nem muito grandes para a gama de valores de $\dot{x}$, x e t considerados.

$$\delta(\dot{x}, x, t) = \frac{-f(\dot{x}, x, t) + \omega^2 x}{\omega^2} \tag{12.27}$$

Usando a Eq. (12.27), podemos escrever a Eq. (12.26) como

$$\ddot{x} + \omega^2 x = \omega^2 \, \delta(\dot{x}, x, t) \tag{12.28}$$

A função $\delta(\dot{x}, x, t)$ depende das variáveis $\dot{x}$, x e t. Para pequenas variações nestas variáveis, entretanto, $\delta(\dot{x}, x, t)$ pode ser considerado constante. Então, nas vizinhanças do estado $x = x_1, \dot{x} = \dot{x}_1, t = t_1$; ou seja, $x = x_1 \pm \Delta x_1$, $\dot{x} = \dot{x}_1 \pm \Delta \dot{x}_1$, $t = t_1 \pm \Delta t_1$, onde $\Delta x_1, \Delta \dot{x}_1, \Delta t_1$ são supostos pequenos, o valor de δ pode ser suposto constante, δ_1, e a Eq. (12.28) pode ser modificada para

$$\ddot{x} + \omega^2 (x - \delta_1) = 0 \tag{12.29}$$

A Eq. (12.29) mostra um movimento harmônico simples. As trajetórias para este sistema são circunferências centradas em $x = \delta_1, \dot{x}/\omega = 0$ no plano de fase normalizado [plano $x - (\dot{x}/\omega)$]. Portanto, podemos ver que para um pequeno incremento nas vizinhanças do estado $x = x_1$, $\dot{x}/\omega = \dot{x}_1/\omega$, $t = t_1$, a trajetória é um arco da circunferência centrada em $x = \delta_1$, $\dot{x}/\omega = 0$ com raio igual a

$$\sqrt{\left(\frac{\dot{x}_1}{\omega}\right)^2 + (x_1 - \delta_1)^2}$$

Nota-se que no método delta é essencial usar um plano de fase normalizado de tal forma que a trajetória da Eq. (12.29) se torna uma circunferência. (Também é necessário usar escalas numéricas iguais ao longo dos eixos x e $\dot{x}/\omega$.) A construção de uma trajetória é vista na Fig. 12.9. O ponto P representa o estado do sistema na

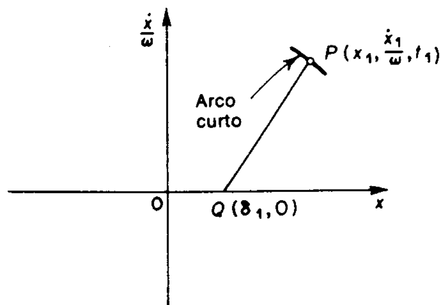


Fig. 12.9 Diagrama mostrando a construção de uma trajetória usando o método delta.

trajetória no instante $t = t_1$. O valor de δ_1 é determinado da Eq. (12.27) como

$$\delta_1 = -\frac{f(\dot{x}_1, x_1, t_1)}{\omega^2} + x_1$$

Uma vez que o ponto Q, o centro do arco de circunferência, está localizado no eixo x, o raio está fixado como sendo $\overline{PQ}$. A trajetória verdadeira nas vizinhanças do ponto P é então aproximada por um pequeno arco de circunferência. O arco deve ser suficientemente pequeno para assegurar que variações nas variáveis são pequenas.

Considere em seguida a equação:

$$\frac{d^2x}{dt^2} + 2\zeta\omega \frac{dx}{dt} + \omega^2 x = 0 \tag{12.30}$$

A Eq. (12.30) contém um termo em x com um coeficiente positivo. Portanto, a Eq. (12.30) pode ser modificada para a forma da Eq. (12.26) simplesmente trazendo o segundo termo da Eq. (12.30) para o lado direito da equação. Ou seja,

$$\frac{d^2x}{dt^2} + \omega^2 x = -2\zeta\omega \frac{dx}{dt} \tag{12.31}$$

A Eq. (12.31) pode ser normalizada substituindo

$$t = \frac{1}{\omega}\tau$$

para fornecer

$$\frac{d^2x}{d\tau^2} + x = -2\zeta \frac{dx}{d\tau} \tag{12.32}$$

Se fizermos

$$\frac{\dot{x}}{\omega} = \frac{dx}{dt}\frac{1}{\omega} = \frac{dx}{d\tau} = y$$

então a Eq. (12.32) se torna

$$\frac{d^2x}{d\tau^2} + x = \delta \tag{12.33}$$

onde $\delta = -2\zeta y$.

Para $y = y_1$ = constante, a Eq. (12.33) indica que a trajetória no plano $x - (dx/d\tau)$ é uma circunferência, cujo centro está localizado em $x = -2\zeta y_1$, $y = 0$. Como a localização do centro depende de y_1, o centro da circunferência se move ao longo do eixo x quando y varia. Ao construir a trajetória da Eq. (12.32), achamos conveniente desenhar uma reta $x = \delta$, ou seja,

$$x = -2\zeta y$$

no plano $x-y$. As Figs. 12.10(a) e (b) mostram a construção de trajetórias.

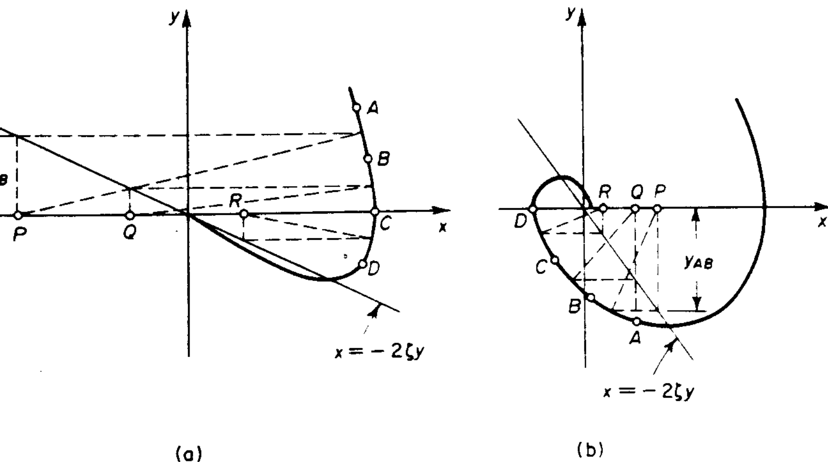


Fig. 12.10 Diagramas mostrando a construção de trajetórias usando o método delta. (a) sistema sobreamortecido; (b) sistema subamortecido.

Em ambas as Figs. 12.10(a) e (b), o centro do arco AB está localizado no ponto P ($x = -2\zeta y_{AB}$, $y = 0$). Aqui y_{AB} é o valor médio de y entre o ponto A e o ponto B. A localização do centro da circunferência pode ser obtida direta e simplesmente usando-se a reta $x = -2\zeta y$ desenhada no plano $x-y$. Nas Figs. 12.10 (a) e (b), os segmentos de arcos do ponto B ao ponto C e do ponto C ao ponto D têm seus centros no ponto Q e ponto R, respectivamente.

O método delta é um método geral, pois as trajetórias podem ser construídas no plano de fase tanto para uma equação de plano de fase representando um sistema físico linear quanto não linear, e, além do mais, independentemente de a equação de plano de fase incluir elementos variantes no tempo ou uma função excitação variando no tempo.

Exemplo 12.6 Considere o sistema descrito pela seguinte equação:

$$\ddot{x} + \dot{x} + x^3 = 0 \tag{12.34}$$

Dadas as condições iniciais $x(0) = 1$, $\dot{x}(0) = 0$, construa a trajetória partindo do ponto inicial. Use o método delta.

Vemos que a Eq. (12.34) não contém um termo em x com um coeficiente positivo. Portanto, é necessário adicionar um termo $\omega^2 x$ em ambos os lados da Eq. (12.34). O valor de ω deve ser escolhido de tal forma que os valores de δ não sejam nem muito pequenos nem muito grandes. Se escolhermos $\omega = 1$, a Eq. (12.34) pode ser escrita como

$$\ddot{x} + x = -\dot{x} - x^3 + x$$
$$\delta = -\dot{x} - x^3 + x$$

Como o valor de δ depende tanto de x como de $\dot{x}$, será necessário fazer tentativas sucessivas no processo de construção.

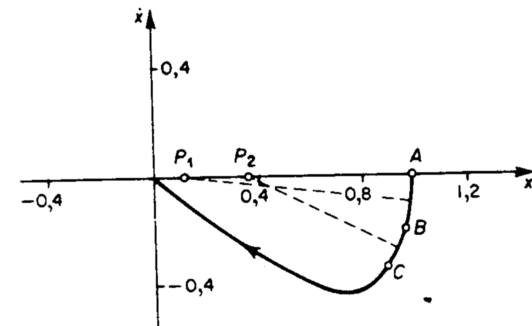


Fig. 12.11 Diagrama de plano de fase mostrando a trajetória partindo do ponto $x(0) = 1$, $\dot{x}(0) = 0$, de um sistema descrito por $\ddot{x} + \dot{x} + x^3 = 0$.

A trajetória começa no ponto A, ($x = 1$, $\dot{x} = 0$), na Fig. 12.11. Nas vizinhanças do ponto A,

$$\delta = -0 - 1 + 1 = 0$$

Portanto, o arco inicial está centrado no ponto (0,0) com raio igual à unidade. Então um arco pequeno é desenhado. O valor médio de x e o valor médio de $\dot{x}$ para este arco são usados para determinar um valor mais preciso de δ. Algumas tentativas sucessivas serão suficientes para obter um valor razoavelmente preciso de δ. Neste exemplo, o primeiro arco AB da Fig. 12.11 está centrado no ponto P_1 ($x = 0{,}12$, $\dot{x} = 0$). Para o segundo arco são feitas tentativas sucessivas similares, mostrando que o segundo arco BC da Fig. 12.11 está centrado no ponto P_2 ($x = 0{,}37$, $\dot{x} = 0$). Continuando desta forma, podemos construir a trajetória até onde desejarmos.

Da discussão que acabamos de fazer, pode-se ver que um sistema com δ, que depende de três variáveis $\dot{x}$, x e t, pode ser tratado da mesma forma. Ao construir trajetórias, entretanto, devemos manter a contagem de t. (Informação temporal pode ser obtida da trajetória. Alguns métodos para se obter informação de um gráfico de plano de fase são dados na seção seguinte.)

12.3 OBTENÇÃO DE SOLUÇÕES TEMPORAIS A PARTIR DE GRÁFICOS DE PLANO DE FASE

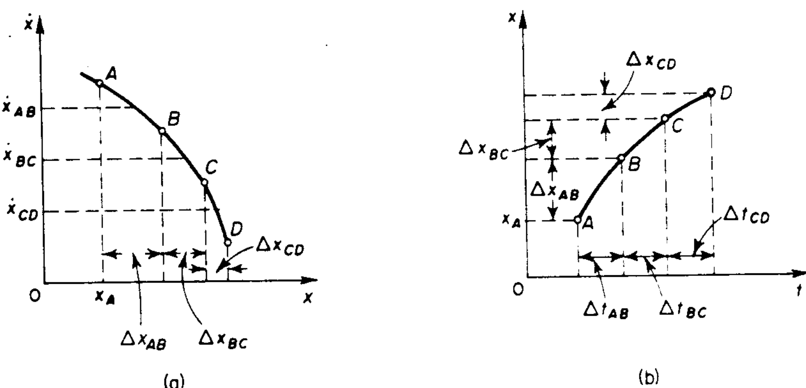
A trajetória desenhada no plano $x-\dot{x}$ é um gráfico de $\dot{x}$ em função de x. O tempo não aparece explicitamente neste gráfico. Para a análise de um sistema, podemos necessitar apenas de trajetórias; entretanto, às vezes é desejável se obter um gráfico da variável x em função do tempo t. Em tal caso, é possível se obter a solução temporal do gráfico do plano de fase, mesmo que a equação diferencial original não possa ser resolvida para x e $\dot{x}$ como funções de t. O processo de obtenção de uma solução temporal é essencialmente passo-a-passo e pode ser feito de várias diferentes maneiras. Uma vez que a trajetória está graduada em t, o comportamento da resposta do sistema com respeito ao tempo pode ser visualizado.

Informação temporal baseada em $\Delta t = \Delta x/\dot{x}$. Pode-se ver que para pequenos incrementos Δx e Δt, a velocidade média é dada por $\dot{x}_{\text{méd.}} = \Delta x/\Delta t$. O tempo

cremental Δt é então

$$\Delta t = \frac{\Delta x}{\dot{x}_{\text{méd.}}} \tag{12.35}$$

A Fig. 12.12(a) mostra uma trajetória no plano $x - \dot{x}$. O tempo incremental Δt_{AB} ecessário para que o ponto representativo atravesse o desvio incremental Δx_{AB} é $t_{AB} = \Delta x_{AB}/\dot{x}_{AB}$. De forma similar, $\Delta t_{BC} = \Delta x_{BC}/\dot{x}_{BC}$. Então a solução temporal $x = t)$ pode ser facilmente desenhada como visto na Fig. 12.12(b). Para boa precisão, desvio incremental Δx deve ser escolhido suficientemente pequeno para que as orrespondentes variações incrementais em $\dot{x}$ e t sejam razoavelmente pequenas. ntretanto, o valor de Δx não precisa ser constante. Ele pode ser mudado de acordo om as formas de porções da trajetória para minimizar o trabalho necessário para se bter uma precisão razoável.



ig. 12.12 (a) Uma trajetória no plano x-$\dot{x}$; (b) um gráfico da solução temporal $x(t)$ em função e t.

Para trajetórias no plano de fase normalizado [plano $x - (\dot{x}/\omega)$], a Eq. (12.35) é modificada para

$$\Delta t = \frac{\Delta x}{\dot{x}_{\text{méd.}}} = \frac{1}{\omega} \frac{\Delta x}{\left(\dfrac{\dot{x}}{\omega}\right)_{\text{méd.}}}$$

Em um caso geral onde as equações diferenciais originais são dadas pelas Eqs. (12.1) e (12.2), Δt pode ser obtida da seguinte relação:

$$\Delta t = \frac{\Delta x_1}{f_1(x_{1\text{méd.}}, x_{2\text{méd.}})} = \frac{\Delta x_2}{f_2(x_{1\text{méd.}}, x_{2\text{méd.}})}$$

onde $x_{1\text{méd.}}$ e $x_{2\text{méd.}}$ são os valores médios de x_1 e x_2 durante um dado intervalo de tempo Δt. Os incrementos em Δx_1, Δx_2, e Δt devem ser razoavelmente pequenos para uma boa precisão.

Informação temporal baseada em $t = \int (1/\dot{x})dx$. Como $\dot{x} = dx/dt$, então o intervalo de tempo $t_2 - t_1$ pode ser expresso como

$$t_2 - t_1 = \int_{x_1}^{x_2} \frac{1}{\dot{x}}\, dx \tag{12.36}$$

A Eq. (12.36) mostra que se a trajetória é refeita com a coordenada $1/\dot{x}$ e a abscissa x, a área sob a curva resultante representa o correspondente intervalo de tempo. O tempo requerido para o ponto representativo na trajetória para ir do ponto A ao ponto B visto na Fig. 12.13(a) pode ser obtido da seguinte equação:

$$t_{AB} = \int_A^B \frac{1}{\dot{x}}\, dx \tag{12.37}$$

O lado direito da Eq. (12.37) é a área hachurada na Fig. 12.13(b). Esta área pode ser obtida analítica ou graficamente por meios convencionais.

Quando o valor de $\dot{x}$ se torna zero entre os limites de integração, o valor de $1/\dot{x}$ se torna infinito, resultando em uma dificuldade no cálculo da integral dada pela Eq. (12.36). Por exemplo, o valor de $\dot{x}$ se torna zero entre o ponto C e o ponto D na Fig. 12.13(a). Portanto, o intervalo de tempo necessário para o ponto representativo se mover do ponto C ao ponto D deve ser computado por alguma outra forma. (Veja o método baseado em aproximações por arcos de circunferência apresentado a seguir.)

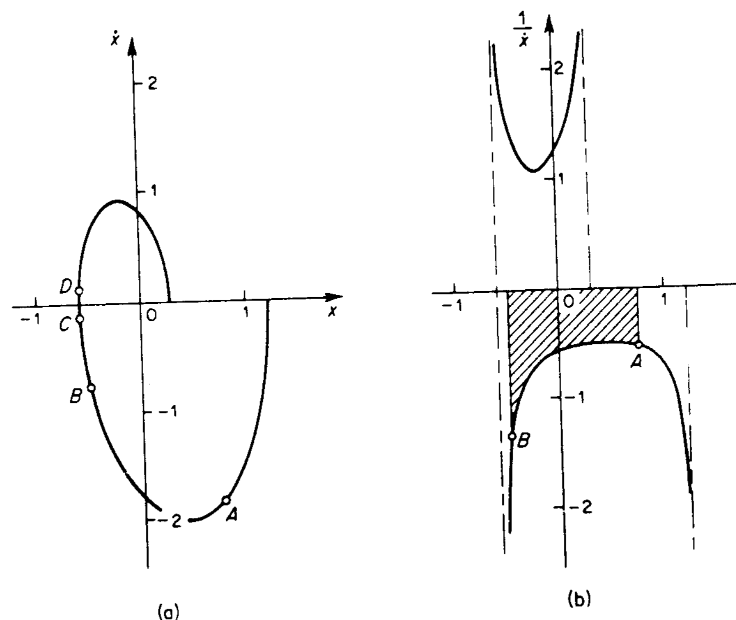


Fig. 12.13 (a) Uma trajetória no plano x-$\dot{x}$; (b) computação gráfica de t_{AB}. (A área hachurada igual a t_{AB}.)

Para trajetórias no plano de fase normalizado [plano $x - (\dot{x}/\omega)$], a Eq. (12.36) é modificada para

$$t_2 - t_1 = \frac{1}{\omega}\int_{x_1}^{x_2} \frac{1}{\left(\frac{\dot{x}}{\omega}\right)} dx$$

Informação temporal baseada em aproximações por arcos de circunferência. Neste método, a trajetória é aproximada por uma seqüência de arcos de circunferência centrados no eixo x. No plano $x - \dot{x}$, se a variável x está em movimento harmônico simples em torno do ponto $(x_0, 0)$ com uma velocidade angular de 1 rad/s, ou seja,

$$\ddot{x} + (x - x_0) = 0$$

então a trajetória pode ser escrita como

$$\dot{x}^2 + (x - x_0)^2 = K^2 \tag{12.38}$$

onde K é uma constante de integração. A Eq. (12.38) indica que a trajetória é uma circunferência de raio K, centrada em $x = x_0$, $\dot{x} = 0$. A velocidade de um ponto representativo ao longo da trajetória é dada por

$$\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{d\dot{x}}{dt}\right)^2}$$

(Esta velocidade não deve ser confundida com a velocidade $\dot{x}$, que é a taxa de variação temporal do deslocamento x.) Em geral, a velocidade de um ponto representativo varia quando ele se move ao longo da trajetória. A velocidade é não nula e finita, exceto em pontos singulares, onde esta velocidade se torna nula, porque tanto dx/dt como $d\dot{x}/dt$ se tornam zero simultaneamente. A velocidade do ponto representativo na trajetória circular é dada por

$$\sqrt{\dot{x}^2 + \ddot{x}^2} = \sqrt{\dot{x}^2 + (x - x_0)^2} = K$$

Isto indica que a velocidade de um ponto representativo ao longo da trajetória é constante. Como o perímetro da circunferência é $2\pi K$ e a velocidade de um ponto representativo é uma constante K, este ponto requer 2π segundos para percorrer uma circunferência completa, portanto o período é 2π segundos.

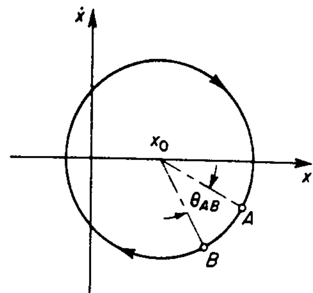


Fig. 12.14 Trajetória circular.

A Fig. 12.14 mostra uma trajetória circular. O tempo necessário para o ponto representativo na trajetória se mover do ponto A ao ponto B pode ser obtido como

$$t_{AB} = 2\pi\frac{\theta_{AB}}{2\pi} = \theta_{AB} \text{ s}$$

onde o ângulo θ_{AB} é a magnitude do ângulo correspondente ao arco circular AB e medido em radianos. Portanto, o tempo t_{AB} em segundos é numericamente igual a θ_{AB} rad se as coordenadas do plano de fase são x e $\dot{x}$. Isto é, se θ_{AB} é igual a 0.5 rad, então t_{AB} é igual a 0.5 s.

Em geral, as trajetórias não são circulares; entretanto, podem ser bem aproximadas por uma série de arcos de circunferências centradas no eixo x. Por exemplo, considere a trajetória vista na Fig. 12.15. Uma parte da trajetória, $ABCD$, pode ser aproximada por três arcos circulares centrados no eixo x. AB pode ser aproximada por uma parte da circunferência centrada no ponto P. Similarmente, BC e CD são partes de circunferências centradas no ponto Q e no ponto R, respectivamente. Então t_{AD}, o tempo necessário para um ponto representativo se mover do ponto A ao ponto D na trajetória, pode ser obtido como

$$t_{AD} = t_{AB} + t_{BC} + t_{CD}$$

onde t_{AB}, t_{BC} e t_{CD} são numericamente iguais a θ_{AB}, θ_{BC} e θ_{CD}, respectivamente; e $\theta_{AB} = \angle APB$, $\theta_{BC} = \angle BQC$ e $\theta_{CD} = \angle CRD$, todos medidos em radianos.

Se o plano de fase é normalizado com x e $\dot{x}/\omega$ como coordenadas, então t_{AB} e θ_{AB} não são numericamente iguais. Vamos considerar este caso com mais detalhes. Se a variável x está em movimento harmônico simples com velocidade angular ω rad/s, ou seja,

$$\ddot{x} + \omega^2(x - x_0) = 0$$

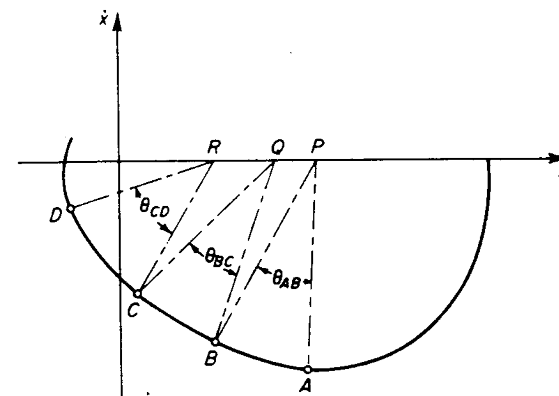


Fig. 12.15 Aproximação de uma trajetória no plano x-$\dot{x}$ por arcos de circunferências.

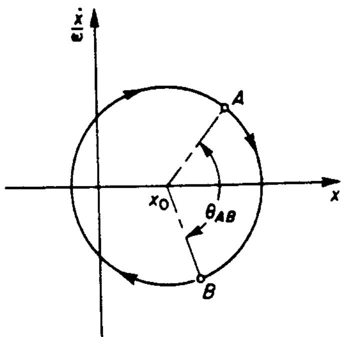


Fig. 12.16 Computação gráfica de t_{AB} no plano $x - (\dot{x}/\omega)$.

a equação da trajetória pode ser escrita como

$$\left(\frac{\dot{x}}{\omega}\right)^2 + (x - x_0)^2 = K^2 \tag{12.39}$$

onde K é uma constante. A trajetória dada pela Eq. (12.39) é uma circunferência centrada no ponto $(x_0, 0)$ no plano $x - (\dot{x}/\omega)$ com raio igual a K. A velocidade de um ponto representativo na trajetória é constante, e este ponto requer $2\pi/\omega$ segundos para percorrer um ciclo completo. Portanto, na Fig. 12.16, t_{AB}, o intervalo de tempo requerido para um ponto representativo se mover do ponto A ao ponto B, pode ser obtido como

$$t_{AB} = \frac{2\pi}{\omega} \times \frac{\theta_{AB}}{2\pi} = \frac{\theta_{AB}}{\omega}\ \text{s}$$

Portanto, no plano $x - (\dot{x}/\omega)$, o intervalo de tempo t_{AB} segundos é igual a θ_{AB}/ω, onde θ_{AB} é medido em radianos e ω é medido em radianos por segundo.

12.4 PONTOS SINGULARES

Nesta seção, examinaremos o comportamento de trajetórias nas vizinhanças de pontos singulares. Considere o sistema

$$\frac{dx_1}{dt} = f_1(x_1, x_2)$$

$$\frac{dx_2}{dt} = f_2(x_1, x_2)$$

onde $f_1(x_1, x_2)$ e $f_2(x_1, x_2)$ são funções analíticas das variáveis x_1 e x_2 nas vizinhanças da origem. Suponha que a origem é um ponto singular, ou um ponto de equilíbrio, de tal forma que

$$f_1(0, 0) = 0,$$

$$f_2(0, 0) = 0$$

Vamos expandir $f_1(x_1, x_2)$ e $f_2(x_1, x_2)$ em série de Taylor nas vizinhanças da origem. O sistema de equações então se torna

$$\frac{dx_1}{dt} = a_1x_1 + b_1x_2 + a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2 + g_1(x_1, x_2) \tag{12.40}$$

$$\frac{dx_2}{dt} = a_2x_1 + b_2x_2 + b_{11}x_1^2 + b_{12}x_1x_2 + b_{22}x_2^2 + g_2(x_1, x_2) \tag{12.41}$$

onde $g_1(x_1, x_2)$ e $g_2(x_1, x_2)$ envolvem apenas potências de ordem três ou superior de x_1 e x_2.

Nas vizinhanças da origem, onde x_1 e x_2 são muito pequenos, as Eqs. (12.40) e (12.41) podem ser aproximadas apenas por termos lineares, contanto que sejam dominantes perto da origem. Então,

$$\frac{dx_1}{dt} = a_1x_1 + b_1x_2 \tag{12.42}$$

$$\frac{dx_2}{dt} = a_2x_1 + b_2x_2 \tag{12.43}$$

Um exame das soluções das Eqs. (12.42) e (12.43) é útil para desenhar as trajetórias perto da origem.

As Eqs. (12.42) e (12.43) podem ser modificadas para

$$\ddot{x} + a\dot{x} + bx = 0 \tag{12.44}$$

onde

$$x = x_1$$
$$a = -a_1 - b_2$$
$$b = a_1b_2 - a_2b_1$$

Note que a Eq. (12.44) é equivalente às Eqs. (12.42) e (12.43). Se as raízes da equação característica da Eq. (12.44) têm partes reais negativas, todas as trajetórias perto da origem se dirigirão a ela quando t cresce indefinidamente. Se pelo menos uma raiz é nula, então a estabilidade não pode ser determinada a partir da equação linearizada, a Eq. (12.44). Neste caso, o comportamento das trajetórias perto da origem depende dos termos de ordem superior nas Eqs. (12.40) e (12.41).

Classificação de pontos singulares. O gráfico de plano de fase de um sistema linear autônomo é uma família de trajetórias que não se cruzam, que descrevem a resposta do sistema a todas as condições iniciais possíveis. Considere a equação diferencial de segunda-ordem linearizada, Eq. (12.44). A localização do ponto singular no plano $x - \dot{x}$ é a origem. A natureza da solução da Eq. (12.44) é determinada pelas duas raízes λ_1 e λ_2 da seguinte equação característica:

$$\lambda^2 + a\lambda + b = 0$$

Aqui supomos que a e b são constantes com $b = 0$. A localização de λ_1 e λ_2 no plano complexo determina as características do ponto singular. Pode-se ver que existem os seguintes seis casos:

1. λ_1 e λ_2 são complexos conjugados e estão no semiplano esquerdo
2. λ_1 e λ_2 são complexos conjugados e estão no semiplano direito
3. λ_1 e λ_2 são reais e estão no semiplano esquerdo
4. λ_1 e λ_2 são reais e estão no semiplano direito.
5. λ_1 e λ_2 são complexos conjugados e estão no eixo $j\omega$.
6. λ_1 e λ_2 são reais, λ_1 está no semiplano esquerdo e λ_2 está no semiplano direito.

De acordo com a natureza das respostas correspondendo a cada caso, os pontos singulares são classificados como foco estável, foco instável, nó estável, nó instável, centro, e ponto de sela, respectivamente. Os gráficos de plano de fase de cada um dos seis casos são ilustrados na Fig. 12.17, e os pontos singulares associados são indicados. Se o ponto singular é um ponto de sela, há trajetórias particulares que entram no ponto de sela e separam o plano de fase em regiões de movimentos distintos. Tais trajetórias são chamadas separatrizes. As trajetórias vistas na Fig. 12.17 podem ser construídas facilmente pelo método das isóclinas. Os gráficos de plano de fase mostram claramente o tipo de resposta do sistema, uma vez que a condição inicial é dada.

A classificação de pontos singulares que acabamos de dar se aplica apenas para sistemas de segunda-ordem. Para sistemas de terceira-ordem ou superior, os pontos singulares são normalmente classificados como estáveis, assintoticamente estáveis, ou instáveis. Para detalhes, referir-se ao Cap. 15.

Ciclos-limite. Ciclos-limite ocorrem freqüentemente em sistemas físicos como osciladores eletrônicos. Ciclos-limite têm uma configuração geométrica distinta no gráfico de plano de fase, ou seja, aquela de um caminho fechado isolado no plano de fase. Um dado sistema pode ter mais que um ciclo-limite. Um ciclo-limite representa uma oscilação estacionária, para a qual ou da qual todas as trajetórias vizinhas irão convergir ou divergir. Portanto ela divide o plano em uma região interna e uma externa. Nenhuma trajetória dentro (fora) de um particular ciclo-limite cruza o ciclo-limite para entrar na região externa (interna).

Um ciclo-limite em um sistema não linear descreve a amplitude e período de uma oscilação mantida. Deve-se ressaltar que nem todas as curvas fechadas no plano de fase são ciclos-limite. Um gráfico de plano de fase de um sistema conservativo, em que não há amortecimento para dissipar energia, é uma família contínua de curvas fechadas. Curvas fechadas desta espécie não são ciclos-limite porque nenhuma destas curvas é isolada da outra. Tais trajetórias sempre ocorrem como uma família contínua, de tal forma que há curvas fechadas em qualquer vizinhança de qualquer curva fechada em particular. Por outro lado, ciclos-limite são movimentos periódicos apresentados apenas por sistemas não lineares não conservativos. (Se o sistema é dissipativo em toda região, então a perda líquida de energia ao longo de qualquer trajetória no plano de fase é positiva, e não pode haver ciclos-limite. Portanto, se um sistema não linear não conservativo apresenta comportamento de ciclo-limite, o amortecimento equivalente deve ser igual a zero.) Como ciclos-limite são isolados um do outro, não pode haver ciclos-limite nas vizinhanças de qualquer um ciclo-limite.

Um ciclo-limite é chamado estável se as trajetórias perto do ciclo-limite, originando-se de dentro ou de fora, convergem para aquele ciclo-limite. Neste caso, o sistema apresenta uma oscilação mantida com amplitude constante. Isto é mostrado na Fig. 12.18(a). A região interna do ciclo-limite é uma região instável no sentido em que as trajetórias divergem para o ciclo-limite, e a região externa é uma

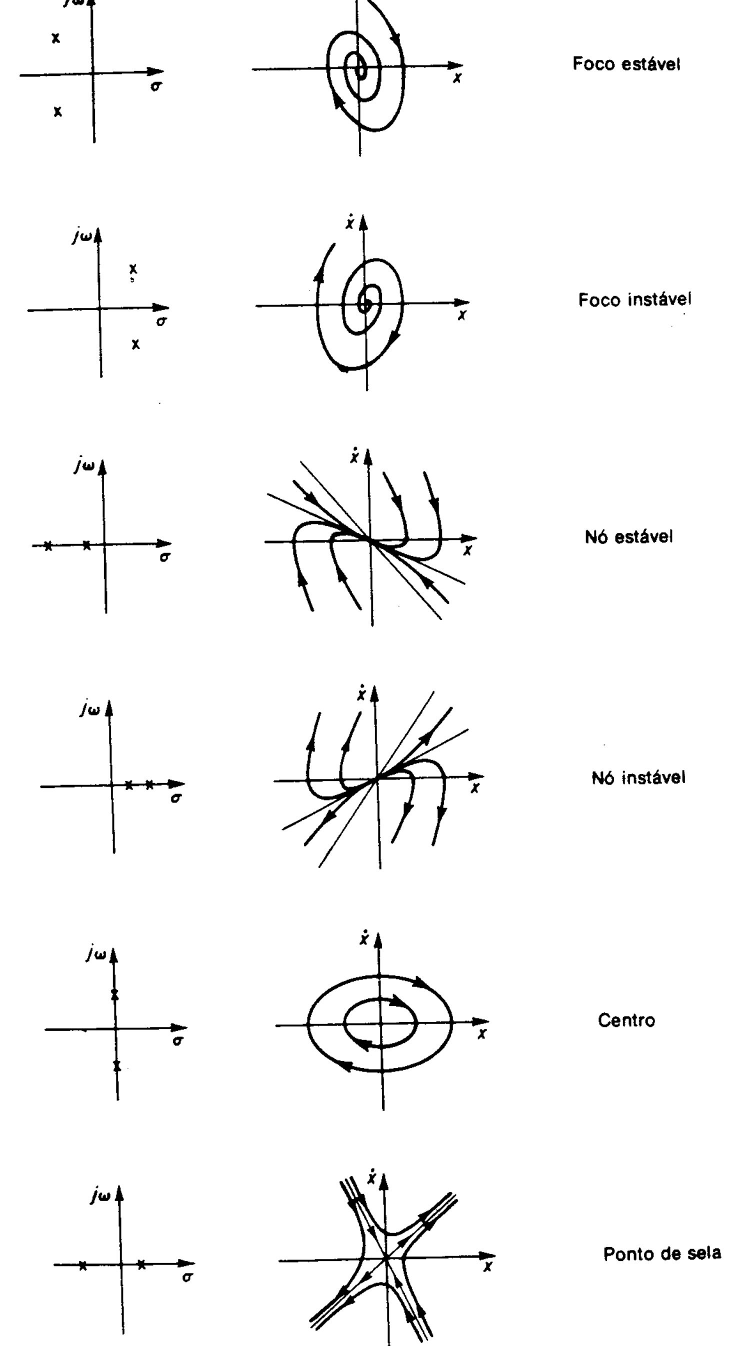


Fig. 12.17 Pontos singulares.

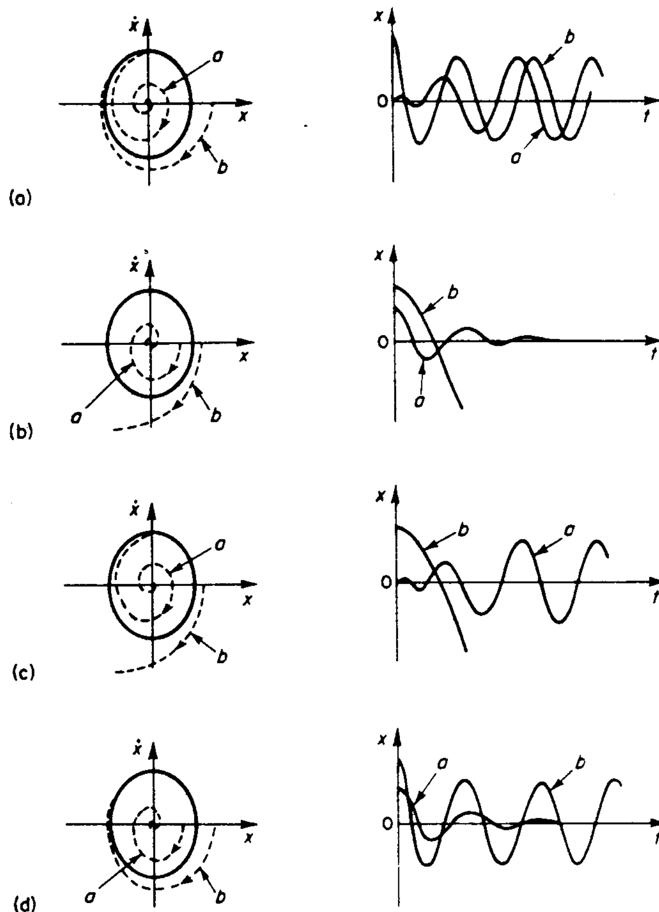


Fig. 12.18 Ciclos-limite e curvas típicas de x em função de t. (a) Ciclo-limite estável, (b) ciclo-limite instável, (c) ciclo-limite semi-estável, (d) ciclo-limite semi-estável.

região estável no sentido em que as trajetórias convergem para o ciclo-limite. No caso de sistemas de controle tendo este tipo de ciclo-limite, o critério de projeto freqüentemente usado é o de fazer a magnitude do ciclo-limite pequena suficiente para satisfazer os requisitos de precisão.

Um ciclo-limite é chamado de instável quando trajetórias próximas divergem deste ciclo-limite. Neste caso, uma região instável circunda uma região estável. Se uma trajetória começa dentro de uma região estável, ela converge para um ponto singular dentro do ciclo-limite. Se uma trajetória começa em uma região instável, entretanto, ela diverge e aumenta para infinito com o tempo. Isto é mostrado na Fig. 12.18(b). A região interna de um ciclo-limite instável é a região estável, e a região externa é a região instável. O critério de projeto para um sistema de controle com um tal ciclo-limite instável é fazer a região estável a maior possível. Operação de

ciclo-limite instável pode ser realizada teoricamente se as condições iniciais podem ser ajustadas com exatidão. Entretanto, qualquer pequena perturbação irá causar instabilidade nesta operação de ciclo-limite.

Um ciclo-limite é chamado semi-estável se as trajetórias que se originam em pontos fora do ciclo-limite divergem deste, enquanto que trajetórias que se originam em pontos dentro do ciclo-limite convergem a este, como visto na Fig. 12.18(c), ou vice-versa, como visto na Fig. 12.18(d). Um sistema de controle pode não ter ciclo-limite, ou então ter um ou mais.

Considere um sistema com dois ciclos-limite e suponha que estes dois ciclos-limite estão localizados perto um do outro. Se o ciclo-limite maior é instável, enquanto que o ciclo-limite menor é estável como visto na Fig. 12.19, então as características do sistema se tornam similares às de um ciclo-limite semi-estável visto na Fig. 12.18(c). De modo similar, se o ciclo-limite maior é estável, enquanto que o menor é instável, as características do sistema se tornam similares à de um ciclo-limite semi-estável, visto na Fig. 12.18(d).

Ciclos-limite estáveis podem ser observados experimentalmente, mas, em virtude de "ruído", o mesmo não se pode dizer de ciclos-limite instáveis e semi-estáveis.

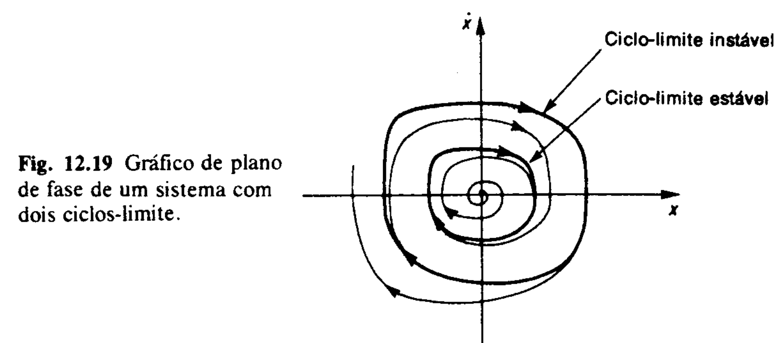


Fig. 12.19 Gráfico de plano de fase de um sistema com dois ciclos-limite.

Com relação à operação de ciclo-limite, devemos notar que, obviamente, deve ser fornecida energia ao sistema de alguma forma, por exemplo, uma tensão DC, vento constante, etc. Em um sistema auto-oscilatório, o trabalho realizado pela fonte de energia, que pode ser constante, é periódico. Em outras palavras, um sistema auto-oscilatório pode gerar um movimento periódico a partir de uma fonte de energia não periódica.

Finalmente, note que exceto em casos simples, é difícil ou impossível determinar a localização exata de ciclos-limite no plano de fase por técnicas analíticas. A localização exata de ciclos-limite pode ser achada apenas por técnicas gráficas, experimentais, ou computacionais.

Exemplo 12.7 Obtenha um gráfico de plano de fase do sistema dado por

$$\ddot{x} + 0{,}5\dot{x} + 2x + x^2 = 0 \tag{12.45}$$

Os pontos singulares para este sistema são

$$x = 0, \dot{x} = 0 \qquad \text{e} \qquad x = -2, \dot{x} = 0$$

A natureza destes pontos pode ser determinada como segue. Em uma vizinhança da origem, a Eq. (12.45) pode ser linearizada para

$$\ddot{x} + 0{,}5\dot{x} + 2x = 0$$

As duas raízes da equação característica

$$\lambda^2 + 0{,}5\lambda + 2 = 0$$

são

$$\lambda_1 = -0{,}25 + j1{,}39, \qquad \lambda_2 = -0{,}25 - j1{,}39$$

Portanto este ponto singular é um foco estável.

Em uma vizinhança do ponto singular $(-2, 0)$, a Eq. (12.45) pode ser escrita como segue: Fazendo

$$y = x + 2$$

obtemos

$$\ddot{y} + 0{,}5\dot{y} - 2y + y^2 = 0 \tag{12.46}$$

Perto do ponto $y = 0, \dot{y} = 0$, a Eq. (12.46) se torna

$$\ddot{y} + 0{,}5\dot{y} - 2y = 0$$

As duas raízes da equação característica

$$\mu^2 + 0{,}5\mu - 2 = 0$$

são

$$\mu_1 = 1{,}19, \qquad \mu_2 = -1{,}69$$

Portanto o ponto singular $(-2, 0)$ é um ponto de sela.

Usando o método das isóclinas, o gráfico de plano de fase, como visto na Fig. 12.20, pode ser obtido. As duas trajetórias que entram no ponto de sela $(-2, 0)$ são separatrizes. Neste sistema uma das separatrizes divide o plano de fase em duas regiões de movimento distinto. Em outras palavras, a solução da Eq. (12.45) tem dois tipos diferentes de trajetórias. O tipo de trajetória depende da condição inicial. Se o ponto inicial está na região hachurada (a região delimitada por uma das duas separatrizes), as trajetórias convergem para a origem. Se o ponto inicial está fora da região hachurada, as trajetórias tendem a infinito. (O desvio crescerá na direção negativa sem limite quando o tempo aumenta indefinidamente.)

12.5 ANÁLISE DE PLANO DE FASE DE SISTEMAS DE CONTROLE LINEARES

O método do plano de fase é bastante útil para analisar sistemas não lineares de segunda-ordem. Antes de apresentarmos a análise de sistemas não lineares, entre-

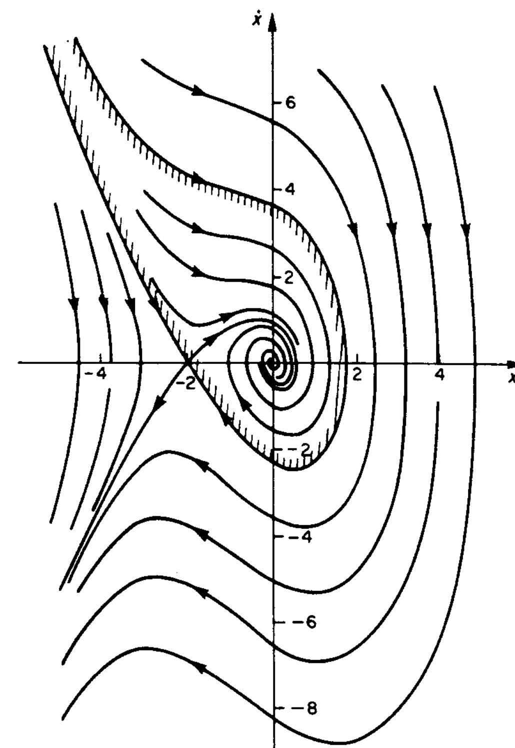


Fig. 12.20 Gráficos de plano de fase de um sistema descrito por $\ddot{x} + 0{,}5\dot{x} + 2x + x^2 = 0$.

tanto, é desejável considerar uma aplicação do método do plano de fase para a análise de sistemas lineares de segunda-ordem porque muitos sistemas de controle com não linearidades dependentes de sinal podem ser aproximados por sistemas lineares por trechos.

Nesta seção, ilustraremos a aplicação deste método na análise de resposta transitória do sistema de controle de segunda-ordem visto na Fig. 12.21.

Resposta a degrau. Suponha que o sistema está inicialmente em repouso. Para este sistema, temos

$$T\ddot{c} + \dot{c} = Ke \tag{12.47}$$

Como $e = r - c$, a Eq. (12.47) pode ser escrita

$$T\ddot{e} + \dot{e} + Ke = T\ddot{r} + \dot{r} \tag{12.48}$$

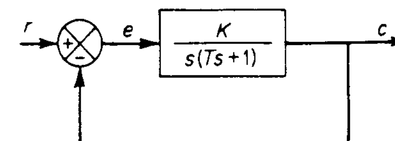


Fig. 12.21 Sistema de controle de segunda-ordem.

Para a entrada em degrau $r(t) = R$, $\ddot{r} = \dot{r} = 0$ para $t > 0$. Portanto, a Eq. (12.48) pode ser escrita como

$$T\ddot{e} + \dot{e} + Ke = 0 \quad \text{para } t > 0 \tag{12.49}$$

Como foi suposto inicialmente que o sistema estava em repouso, as condições iniciais para o sinal de erro são $e(0) = R$ e $\dot{e}(0) = 0$. A trajetória no plano $e - \dot{e}$ começa no ponto $(R, 0)$ e converge para a origem, o ponto singular do sistema. A Fig. 12.22(a) mostra a trajetória quando as raízes da equação característica do sistema são um par complexo conjugado e estão no semiplano esquerdo. A Fig. 12.22(b) mostra a trajetória quando as raízes são reais e estão no semiplano esquerdo. Em ambos os casos, o erro do sistema é nulo em regime estacionário. A natureza da resposta pode ser claramente vista nos gráficos de plano de fase. Por exemplo, o sobre-sinal máximo no caso subamortecido pode ser obtido do gráfico de plano de fase da Fig. 12.22(a).

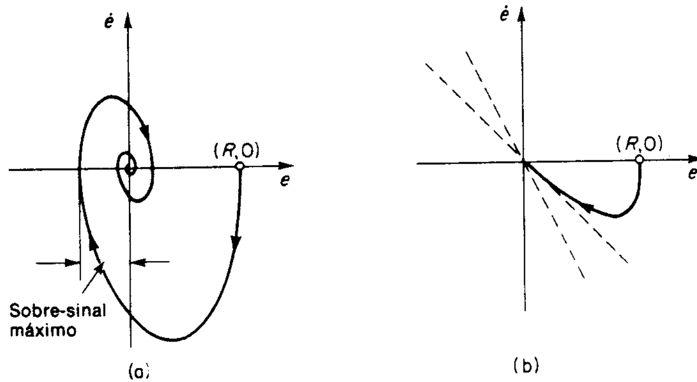


Fig. 12.22 Trajetórias correspondendo a respostas a degrau do sistema da Fig. 12.21. (a) Caso subamortecido, (b) caso sobreamortecido.

Resposta em rampa. Para a entrada em rampa $r(t) = Vt$ ou rampa mais degrau $r(t) = Vt + R$, as derivadas de $r(t)$ se tornam $\ddot{r} = 0$ e $\dot{r} = V$ para $t > 0$. Portanto, a Eq. (12.48) pode ser escrita como

$$T\ddot{e} + \dot{e} + Ke = V \tag{12.50}$$

Seja

$$e - \frac{V}{K} = x$$

Então a Eq. (12.50) se torna

$$T\ddot{x} + \dot{x} + Kx = 0 \tag{12.51}$$

O gráfico de plano de fase da equação diferencial dada pela Eq. (12.51) no plano $x - \dot{x}$ é o mesmo que o da Eq. (12.49) no plano $e - \dot{e}$.

Nota-se que a natureza do ponto singular é determinada pelas raízes da equação característica. A localização do ponto singular no plano $e - \dot{e}$, em geral, depende tanto dos componentes em degrau como da rampa da entrada. (Neste exemplo em particular, a localização do ponto singular no plano de fase depende apenas da entrada em rampa porque o lado direito da Eq. (12.48) não contém um termo em r.) Como é suposto que o sistema está inicialmente em repouso, as condições iniciais para o sinal de erro podem ser escritas como

$$e(0) = R, \qquad \dot{e}(0) = V$$

onde R pode ser zero. As Figs. 12.23(a) e (b) mostram trajetórias no plano $e - \dot{e}$. [A Fig. 12.23(a) corresponde ao caso em que as raízes da equação característica, correspondendo à Eq. (12.49), são um par complexo-conjugado e estão no semiplano esquerdo. A Fig. 12.23(b) corresponde ao caso em que as raízes são reais e estão no semiplano esquerdo.] Nas Figs. 12.23(a) e (b), para entrada em rampa a trajetória começa, por exemplo, no ponto A. Para uma entrada em rampa mais degrau a trajetória começa, por exemplo, no ponto B. Em ambos os casos, as trajetórias convergem para o ponto singular $(V/K, 0)$.

Da análise feita, pode-se ver que o gráfico de plano de fase do sistema visto na Fig. 12.21 para uma entrada em rampa ou rampa mais degrau é o mesmo que aquele para uma entrada em degrau, exceto que todo o gráfico está deslocado para a direita de V/K. O erro estacionário da saída do sistema, quando o sistema está sujeito à entrada $r(t) = Vt + R$, é V/K.

Note que para um sistema com um ponto singular é possível deslocar a posição do ponto singular para a origem de um novo plano de fase introduzindo-se uma nova variável. Portanto, para tal sistema, as características de resposta a degrau e resposta a rampa podem ser mostradas em um único gráfico de plano de fase. Para um sistema não linear com dois ou mais pontos singulares, entretanto, é impossível deslocar todos estes pontos singulares para a origem ao mesmo tempo. Portanto, para um tal sistema, é imperativo se usar um gráfico de plano de fase separado para uma resposta a degrau e para uma resposta a rampa.

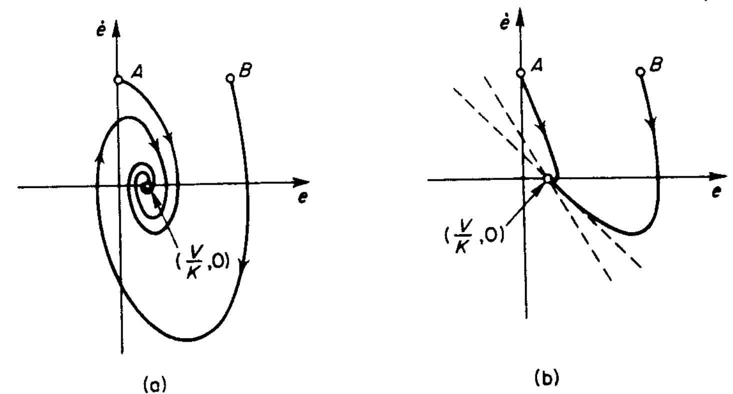


Fig. 12.23 Trajetórias correspondendo a respostas à rampa do sistema da Fig. 12.21. (a) Caso subamortecido, (b) caso sobreamortecido.

Resposta impulsiva. Para uma entrada impulso unitário, a equação do sistema é

$$T\ddot{c} + \dot{c} + Kc = 0 \qquad \text{para } t > 0$$

com condições iniciais $c(0-) = \dot{c}(0-) = 0$ e

$$c(0+) = \lim_{s \to \infty} \frac{sK}{Ts^2 + s + K} = 0, \qquad \dot{c}(0+) = \lim_{s \to \infty} \frac{s^2K}{Ts^2 + s + K} = \frac{K}{T}$$

O ponto de partida da trajetória no plano $c - \dot{c}$ é $(0, K/T)$. Em termos do sinal de erro,

$$T\ddot{e} + \dot{e} + Ke = 0 \qquad \text{para } t > 0$$

com condições iniciais

$$e(0+) = 0, \qquad \dot{e}(0+) = -\frac{K}{T}$$

No plano $e - \dot{e}$, o ponto de partida (correspondendo a $t = 0+$) da trajetória é $(0, -K/T)$. As trajetórias indicando respostas a impulso unitário para $t > 0$ são vistas nas Figs. 12.24 e 12.25. A Fig. 12.24 corresponde a um sistema subamortecido, e a Fig. 12.25 corresponde a um sistema sobreamortecido. Nestas figuras, os pontos de partida (em $t = 0+$) das trajetórias são mostrados pelo ponto P.

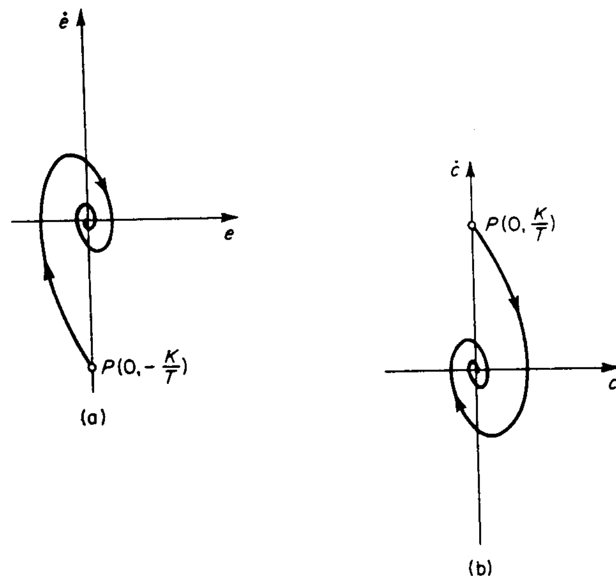


Fig. 12.24 Trajetórias correspondendo a respostas a impulsos unitários do sistema da Fig. 12.21 (caso subamortecido). (a) Gráfico no plano e-$\dot{e}$, (b) gráfico no plano c-$\dot{c}$.

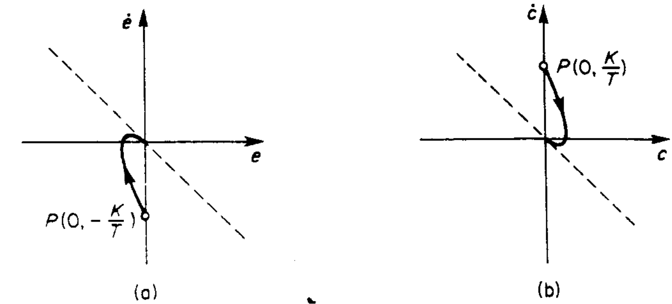


Fig. 12.25 Trajetórias correspondendo a respostas a impulso unitário do sistema da Fig. 12.21 (caso sobreamortecido). (a) Gráfico no plano e-$\dot{e}$, (b) gráfico no plano c-$\dot{c}$.

12.6 ANÁLISE DE PLANO DE FASE DE SISTEMAS DE CONTROLE NÃO LINEARES*

Em sistemas de segunda-ordem com não linearidades dependentes de sinal, é possível aproximar o sistema por vários sistemas lineares por trechos. O plano de fase inteiro é dividido em várias sub-regiões, cada uma correspondendo a uma operação linear individual. Há um ponto singular para cada região, embora ele possa estar localizado fora daquela região particular.

Se o ponto singular está dentro de sua sub-região, então ele é chamado de um ponto singular real. Mas, se o ponto singular está fora da sub-região a que ele pertence e, portanto, nunca pode ser alcançado pelas suas próprias trajetórias, ele é chamado de um ponto singular *virtual*. Um sistema de segunda-ordem com não linearidade dependente de sinal pode ter apenas um ponto singular real.

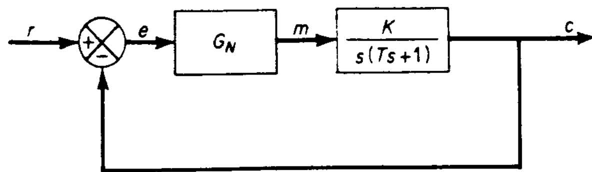
Todas as regiões adjacentes a uma região com um ponto singular real terão pontos singulares virtuais. A localização e natureza de cada ponto singular são determinadas a partir da equação diferencial que governa a dada sub-região. A localização de um ponto singular pode depender da entrada. O gráfico de plano de fase de cada sub-região é o de um sistema linear. A trajetória composta, que é obtida juntando trajetórias nas fronteiras de cada região de operação, fornece a resposta transitória do sistema não linear.

Sistemas de controle com ganhos não lineares. Considere o sistema não linear visto na Fig. 12.26(a). O bloco chamado G_N é o elemento de ganho não linear. A curva característica de entrada-saída deste elemento é vista na Fig. 12.26(b). O ganho do elemento é unitário ou k, sempre que a magnitude do sinal de erro e é maior do que ou menor do que e_0, respectivamente. Ou seja,

$$m = e \qquad \text{para } |e| > e_0 \tag{12.52}$$

$$m = ke \qquad \text{para } |e| < e_0 \tag{12.53}$$

*Esta seção segue a Referência K-1.



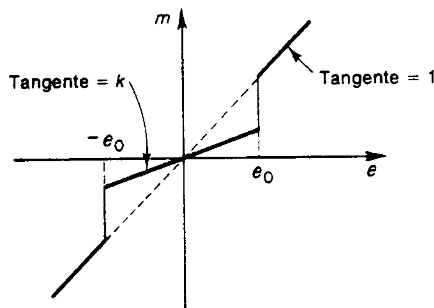


Fig. 12.26 (a) Sistema não linear, (b) curva característica de entrada-saída do elemento de ganho não linear.

O sistema tem um ganho grande para sinais de erro grandes e tem um ganho pequeno para sinais de erro pequenos. O chaveamento do ganho pode ser conseguido usando-se um dispositivo de chaveamento que muda o ganho do amplificador abruptamente de um valor para outro. Para valores pequenos de e, o sistema apresenta uma resposta lenta, e para grandes erros, uma resposta rápida. Esta característica pode ser desejável para sistemas sujeitos a ruídos de baixa amplitude e alta freqüência uma vez que sinais de ruído indesejáveis serão suprimidos substancialmente, enquanto que sinais de comando podem ser transmitidos de modo satisfatório.

Como este sistema é linear por trechos, é possível obter trajetórias para este sistema combinando as trajetórias de dois sistemas lineares.

Neste exemplo, supomos que o sistema está inicialmente em repouso. A equação diferencial relacionando as variáveis c e m é

$$T\ddot{c} + \dot{c} = Km$$

Como $e = r - c$, esta última equação pode ser reescrita como

$$T\ddot{e} + \dot{e} + Km = T\ddot{r} + \dot{r} \qquad (12.54)$$

Dois tipos de entradas serão consideradas aqui, a entrada em degrau e a entrada em rampa (ou rampa mais degrau).

Para entradas em degrau. Como para entradas em degrau $\ddot{r} = \dot{r} = 0$ para $t > 0$, a Eq. (12.54) se torna

$$T\ddot{e} + \dot{e} + Km = 0 \quad \text{para } t > 0 \qquad (12.55)$$

Das Eqs. (12.52), (12.53) e (12.55), obtemos as seguintes duas equações:

$$T\ddot{e} + \dot{e} + Ke = 0 \quad \text{para} |e| > e_0 \qquad (12.56)$$

$$T\ddot{e} + \dot{e} + kKe = 0 \quad \text{para} |e| < e_0 \qquad (12.57)$$

Para se obter a trajetória do sinal de erro, os pontos singulares do sistema devem ser inicialmente determinados. Os pontos singulares estão onde $\dot{e} = \ddot{e} = 0$. Está claro, a partir das Eqs. (12.56) e (12.57), que a origem (0, 0) é o ponto singular para o sistema representado por estas duas equações. Na análise que se segue, será suposto que as raízes da equação característica, correspondendo à Eq. (12.56), são complexos conjugados e estão no semiplano esquerdo do plano complexo. Portanto, a natureza do ponto singular (0, 0) neste caso é aquela de um foco estável.

O valor de k é suposto tal que o coeficiente de amortecimento da equação característica correspondendo à Eq. (12.57) é igual a um. Então, a resposta é subamortecida para erros grandes e criticamente amortecida para erros pequenos. Se a relação entre m e e fosse dada ou por $m = ke$ ou $m = e$ para todas as magnitudes do sinal de erro, então o gráfico de plano de fase teria o aspecto visto na Fig. 12.27(a) ou aquele visto na Fig. 12.27(b).

A Fig. 12.28 mostra a trajetória do sinal de erro do sistema visto na Fig. 12.26(a), com as constantes do sistema sendo $T = 1$, $K = 4$, $k = 0{,}0625$, e $e_0 = 0{,}2$. Supõe-se que a entrada do sistema é um degrau unitário. Como visto na Fig. 12.28, o plano de fase está dividido em três regiões. Na região limitada pelas linhas $e = e_0$ e $e = -e_0$, ocorre uma operação linear representada pela Eq. (12.57). Fora desta

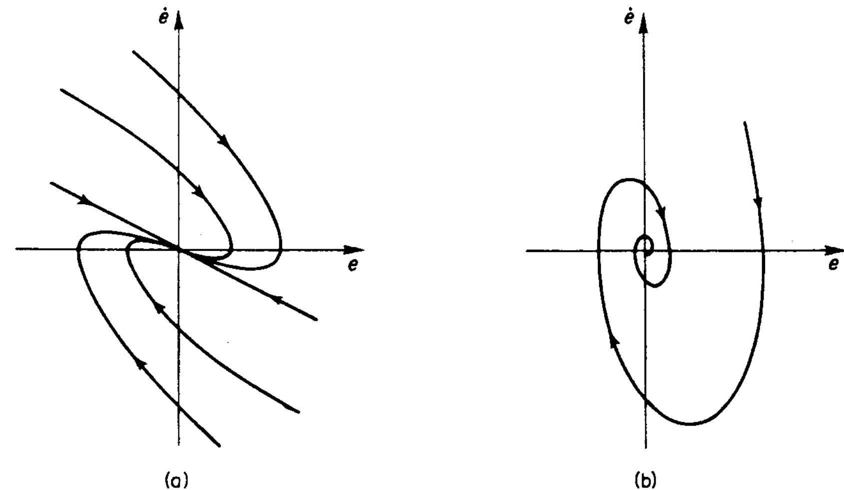


Fig. 12.27 Gráficos de plano de fase de sistemas não lineares. (a) Caso sobreamortecido ($m = ke$), (b) caso subamortecido ($m = e$).

região. ocorrerá a outra operação linear, correspondendo à Eq. (12.56). Na Fig. 12.28, a trajetória começando no ponto A, que é determinada pelas condições iniciais $e(0) = 1, \dot{e}(0) = 0$, tende a convergir para o foco estável (0, 0). A operação do sistema é chaveada no ponto B, entretanto, onde a trajetória intercepta a linha de fronteira $e = e_0$. Neste ponto, a trajetória é conectada a uma trajetória pertencendo ao ponto singular da região adjacente e passando por este ponto. A partir do ponto B, até que a operação do sistema seja novamente chaveada, a trajetória tende a convergir para o nó estável (0, 0). A operação do sistema muda novamente no ponto C, e a trajetória tende a convergir para o foco estável (0, 0) até que o ponto D é alcançado. Neste ponto, a operação do sistema é chaveada mais uma vez. Repetindo o mesmo processo, a trajetória finalmente converge para o nó estável (0, 0). Em regime estacionário, não há erro.

Pode-se ver que a trajetória vista na Fig. 12.28 representa de forma geral características de resposta a degrau mais desejáveis do que aquelas trajetórias vistas nas Figs. 12.27(a) e (b), no sentido de que a primeira resposta é mais rápida que a última. Para pequenas entradas em degrau, a resposta não exibe sobre-sinais. Para entradas em degrau de amplitude média, a resposta do sistema exibe um único sobre-sinal. Para entradas em degraus maiores, pode haver sobre-sinais e "subsinais" *(undershoot)* nas curvas de resposta. A Fig. 12.29 mostra curvas típicas de resposta a degrau.

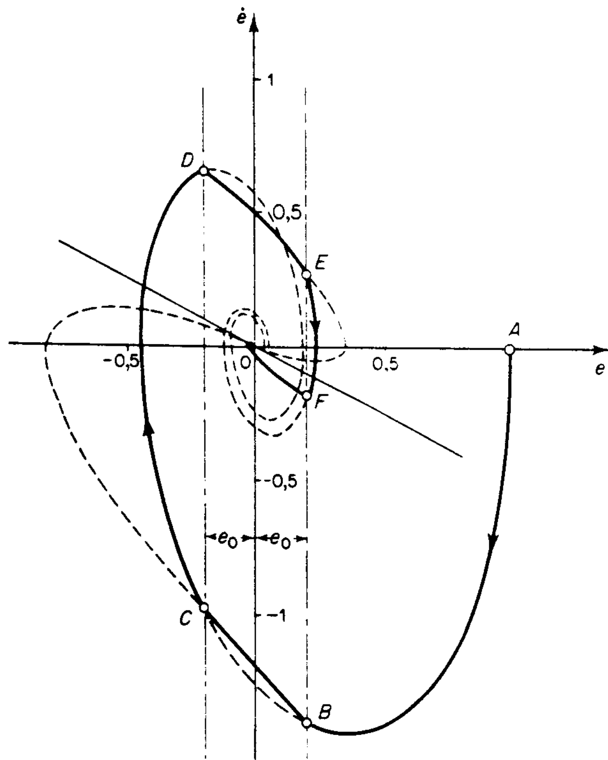


Fig. 12.28 Trajetória correspondendo a uma resposta a degrau unitário do sistema visto na Fig. 12.26(a).

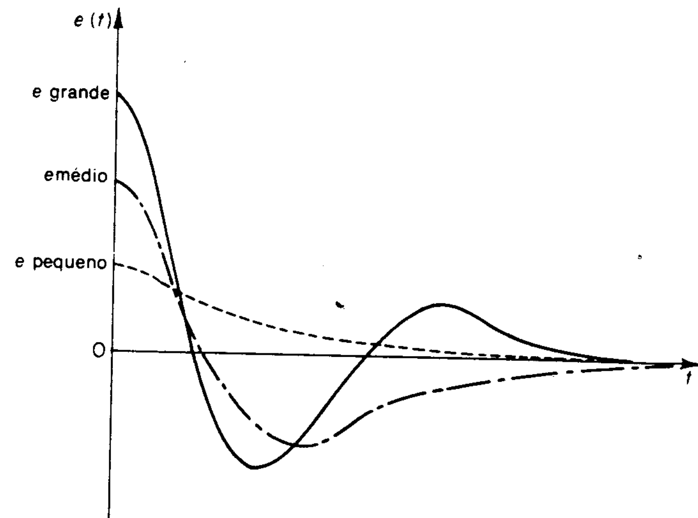


Fig. 12.29 Curvas típicas de resposta a degrau para o sistema visto na Fig. 12.26(a).

Para entradas em rampa (ou rampa mais degrau): Para o sinal de entrada

$$r(t) = R + Vt$$

a Eq. (12.54) se torna

$$T\ddot{e} + \dot{e} + Km = V \quad \text{para } t > 0 \tag{12.58}$$

Das Eqs. (12.52), (12.53) e (12.58), obtemos para $t > 0$

$$T\ddot{e} + \dot{e} + Ke = V \quad \text{para } |e| > e_0 \tag{12.59}$$

$$T\ddot{e} + \dot{e} + Kke = V \quad \text{para } |e| < e_0 \tag{12.60}$$

O ponto singular correspondente à Eq. (12.59) é $(V/K, 0)$, que é suposto ser um foco estável, e o correspondente à Eq. (12.60) é $(V/Kk, 0)$, que se acredita ser um nó estável.

A localização dos pontos singulares pode também ser obtida graficamente. Os valores de e, correspondendo às intersecções da reta horizontal $m = V/K$ com a reta $m = ke$ e com a reta $m = e$, fornecem os valores de e dos pontos singulares. Isto é mostrado nas Figs. 12.30(a), (b) e (c). Para valores pequenos de V, ou seja, para $|V| < kKe_0$, os pontos singulares estão localizados dentro da faixa delimitada pelas retas $e = \pm e_0$ no plano $e - \dot{e}$. Para valores médios de V, ou seja, $kK_0e_0 < |V| < Ke_0$, o ponto singular, correspondendo à Eq. (12.59), está dentro da faixa delimitada pelas retas $e = \pm e_0$ no plano $e - \dot{e}$, e o outro ponto singular, correspondendo à Eq. (12.60), está fora desta faixa. Para valores maiores que V, ou seja, $|V| > Ke_0$, os pontos singulares estão fora desta faixa.

Se a relação entre m e e fosse tal que $m = ke$ durante a resposta temporal, então o gráfico de plano de fase do sistema sujeito à entrada $r(t) = R + Vt$ seria o mesmo

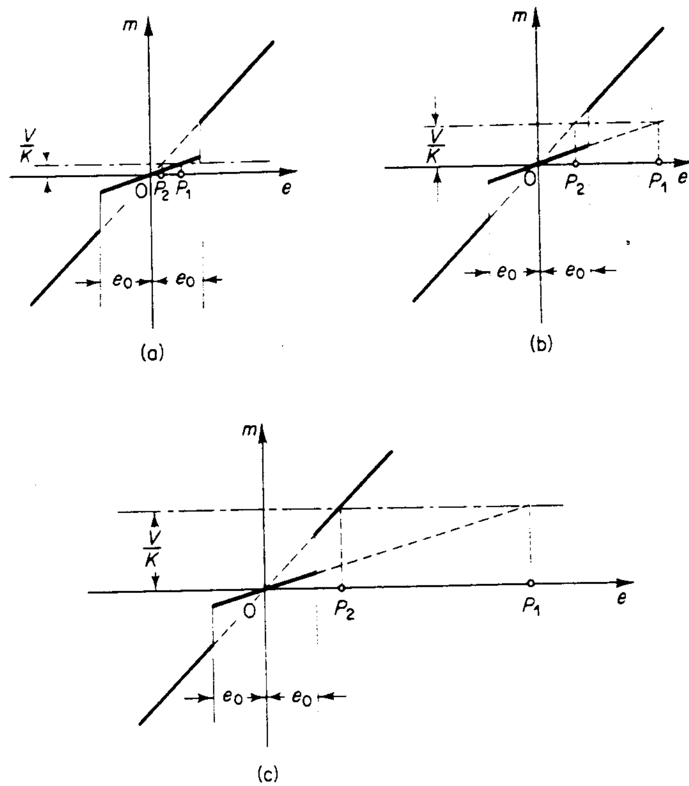


Fig. 12.30 Diagramas mostrando as localizações de pontos singulares.

que aquele visto na Fig. 12.27(a), com exceção de que o primeiro está deslocado para a direita de um valor V/Kk, como visto na Fig. 12.31(a). Similarmente, se m fosse igual a e durante a resposta temporal, então o gráfico do plano de fase seria o mesmo que aquele visto na Fig. 12.27(b), exceto que o primeiro estaria deslocado para a direita pela quantidade V/K, como visto na Fig. 12.31(b).

A Fig. 12.32 mostra a trajetória do sinal de erro para o caso $V < kKe_0$. (Os valores numéricos usados para este exemplo são $T = 1$, $K = 4$, $k = 0{,}0625$, $e_0 = 0{,}2$, $R = 0{,}3$, e $V = 0{,}04$.) O ponto de partida A da trajetória é determinado pelas condições iniciais $e(0) = R = 0{,}3$ e $\dot{e}(0) = V = 0{,}04$. A trajetória partindo do ponto A tende a convergir para o foco estável P_2. Tão logo a trajetória alcance o ponto B, entretanto, a operação do sistema é chaveada, e a trajetória começa a convergir para o nó estável P_1. Em regime permanente, o sistema tem um erro, cuja magnitude é $\overline{OP}_1$.

A Fig. 12.33 mostra a trajetória do sinal de erro quando $kKe_0 < V < Ke_0$. (Na Fig. 12.33, as constantes do sistema são $T = 1$, $K = 4$, $k = 0{,}0625$, e $e_0 = 0{,}2$. A entrada é $r(t) = Vt = 0{,}4t$.) A trajetória começa no ponto A, que corresponde às condições iniciais $e(0) = 0$, $\dot{e}(0) = 0{,}4$, e tende a convergir para o nó estável P_1, (1,6, 0). O chaveamento na operação ocorre no ponto B, e então a trajetória tende a convergir para o foco estável P_2 (0.1, 0). O sistema novamente chaveia no ponto C, e

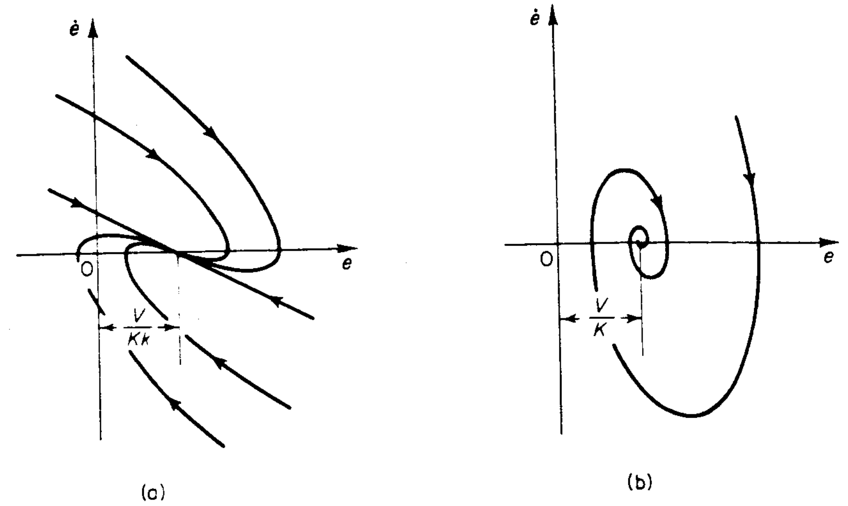


Fig. 12.31 Gráficos de plano de fase de sistemas não lineares com entradas em rampa. (a) Caso sobreamortecido ($m = ke$), (b) caso subamortecido ($m = e$).

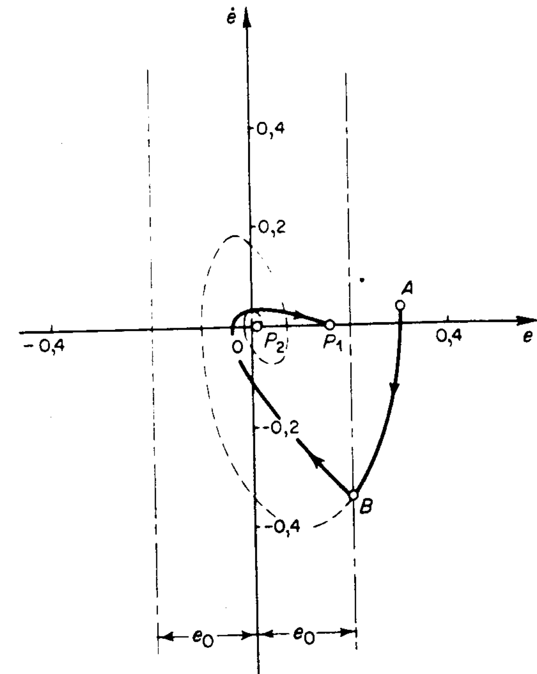


Fig. 12.32 Trajetória correspondendo a uma resposta a rampa (rampa + degrau) do sistema visto na Fig. 12.26(a).

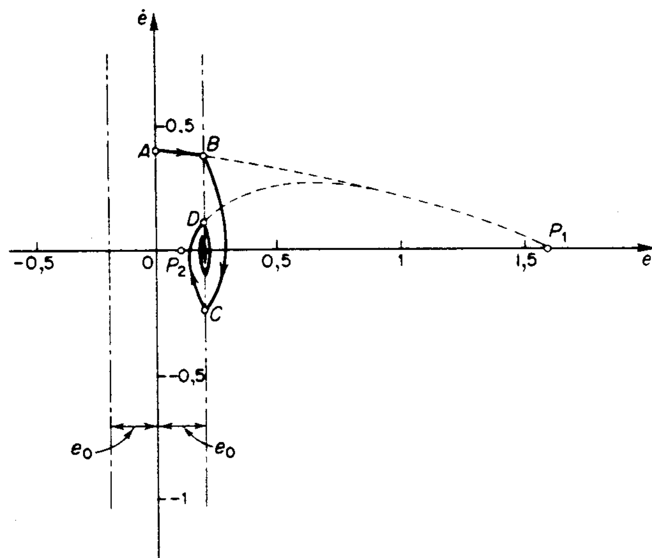


Fig. 12.33 Trajetória correspondendo a uma resposta a rampa do sistema da Fig. 12.26(a).

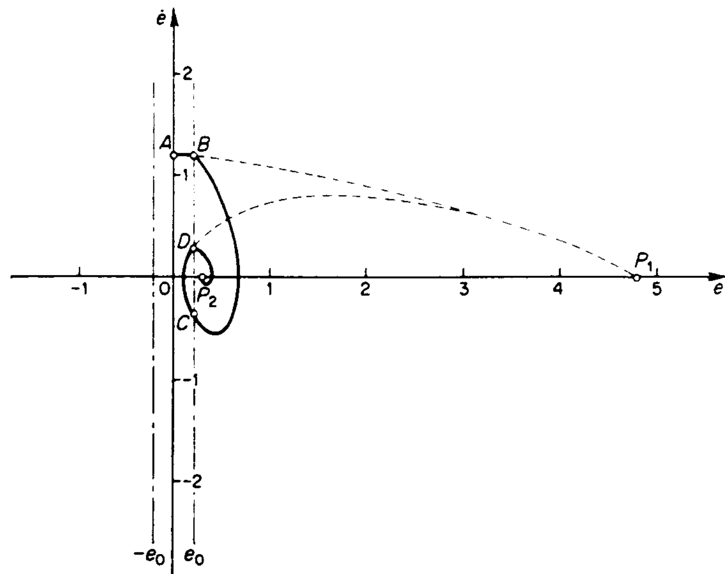


Fig. 12.34 Trajetória correspondendo a uma resposta a rampa do sistema da Fig. 12.26(a).

a trajetória tende a convergir para o nó estável P_1. Um outro chaveamento ocorre na operação do sistema no ponto D. O mesmo processo continua até que a trajetória converge para o ponto onde $e = e_0$, $\dot{e} = 0$. Pode-se ver, da Fig. 12.33, que à medida que a trajetória se aproxima do ponto singular $(e_0, 0)$, o sinal de erro exibe pequenas oscilações, e a magnitude do erro estacionário se torna e_0.

A operação estacionária do sistema depende da construção física do elemento de ganho não linear. Se o chaveamento de uma operação linear para outra envolve algum atraso, a resposta do sistema apresentará um ciclo limite em torno do ponto $(e_0, 0)$. Se o chaveamento ocorre instantaneamente, o sistema, em regime estacionário, apresenta uma espécie de oscilação tudo-ou-nada (em inglês, *chattering*).

A Fig. 12.34 mostra a trajetória do sinal de erro para o caso em que $V > Ke_0$. [A entrada é $r(t) = Vt = 1{,}2t$.] A trajetória partindo do ponto A, com o ponto correspondendo à condição inicial, tende a convergir para o foco estável P_1, (4,8, 0). Entretanto, a operação do sistema é chaveada no ponto B, e a trajetória tende a convergir para o foco estável P_2, (0,3, 0). A operação é novamente chaveada no ponto C, e depois no ponto D, e finalmente a trajetória converge para o foco estável P_2. À medida que a trajetória se aproxima da vizinhança do ponto P_2, o sinal de erro apresenta pequenas oscilações, que irão decair para zero. Em regime estacionário, a magnitude do erro é $\overline{OP}_2$.

A análise feita até agora ilustrou que os tipos de resposta de sistemas não lineares são dependentes da entrada. Por exemplo, para uma pequena entrada em degrau, a resposta não apresenta sobre-sinal. Para uma entrada em degrau grande, a resposta é oscilatória. Para uma entrada em rampa de pequena amplitude, a resposta é aperiódica, enquanto que para uma entrada em rampa de amplitude grande a resposta é oscilatória.

Exemplo 12.8 Considere o sistema visto na Fig. 12.35(a). A curva característica do elemento não linear é vista na Fig. 12.35(b). Supõe-se que não há atraso de tempo no chaveamento da

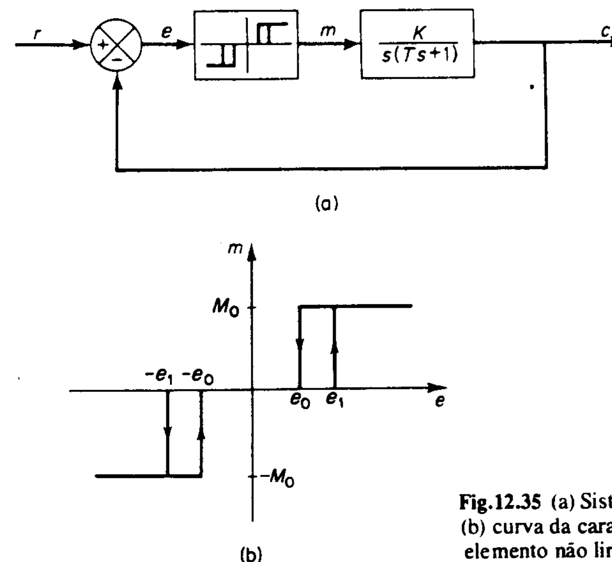


Fig.12.35 (a) Sistema de controle não linear; (b) curva da característica entrada-saída do elemento não linear.

condição ligado para a condição desligado. Determine o comportamento de resposta a degrau e a rampa para este sistema. Supõe-se que as constantes do sistema são $T = 1$, $K = 4$, $e_0 = 0{,}1$, $e_1 = 0{,}2$, e $M_0 = 0{,}2$.

A equação para este sistema é

$$T\ddot{e} + \dot{e} + Km = T\ddot{r} + \dot{r}$$

Para $\dot{e} > 0$, temos

$$m = M_0 \quad \text{para } e > e_1$$
$$m = 0 \quad \text{para } e_1 > e > -e_0$$
$$m = -M_0 \quad \text{para } e < -e_0$$

Para $\dot{e} < 0$, temos

$$m = M_0 \quad \text{para } e > e_0$$
$$m = 0 \quad \text{para } e_0 > e > -e_1$$
$$m = -M \quad \text{para } e < -e_1$$

Para entradas em degrau: A trajetória deste sistema com uma entrada degrau unitário é vista na Fig. 12.36. A resposta mostra um ciclo limite em regime estacionário. Portanto, a oscilação na saída do sistema continua indefinidamente.

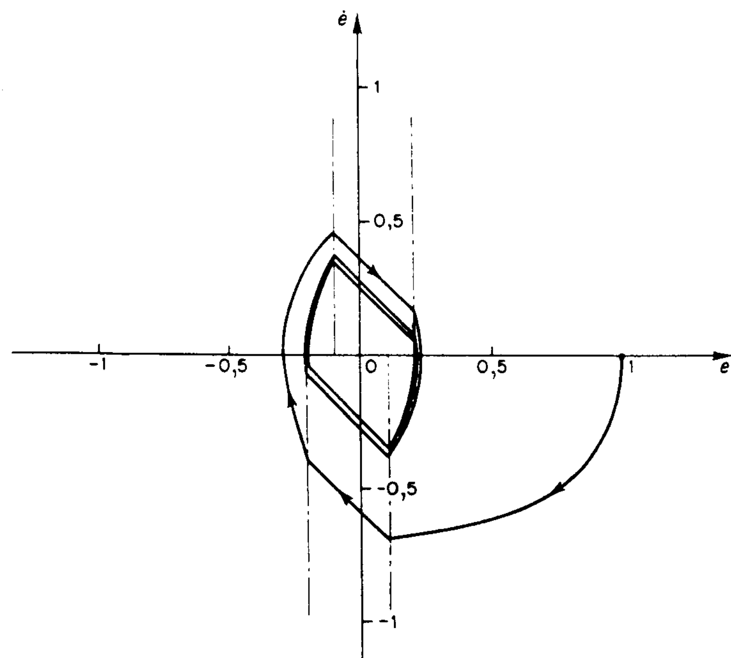


Fig. 12.36 Trajetória correspondendo a uma resposta a degrau unitário do sistema visto na Fig. 12.35(a).

Para entradas em rampa: Para uma entrada em rampa $r(t) = Vt$, a equação do sistema se torna

$$T\ddot{e} + \dot{e} + Km = V$$

Investigaremos agora a resposta para a entrada em rampa para três casos diferentes.
Caso 1 ($V > KM_0$): Um gráfico de plano de fase para o caso em que $V = 1{,}2$ é visto na Fig. 12.37. A trajetória partindo do ponto A, por exemplo, seguirá o caminho $ABCD$. A trajetória é assintótica à reta $\dot{e} = 0{,}4$. O erro tende a infinito quando o tempo aumenta indefinidamente.
Caso 2 ($V < KM_0$): Um gráfico de plano de fase para o caso em que $V = 0{,}4$ é visto na Fig. 12.38. A trajetória partindo do ponto A segue o caminho $ABCDEF$ e converge para um ciclo limite. A oscilação de saída continua indefinidamente.
Caso 3 ($V = KM_0$): Um gráfico de plano de fase para o caso em que $V = 0{,}8$ é visto na Fig. 12.39. A trajetória partindo do ponto A segue o caminho $ABCD$. A trajetória converge para o ponto D. Em regime estacionário, o sistema terá um erro estacionário igual a $\overline{OD}$.

Resumo. A matéria apresentada nesta seção pode ser resumida como segue: Considere o sistema visto na Fig. 12.40. O bloco N no elo direto representa um elemento de ganho não linear que é dependente da amplitude. Se $n(e)$ assume valores univocamente, é uma função contínua e diferenciável de e, então sua inclinação $dn(e)/de = n'(e)$ define um ganho incremental de malha. Enquanto a trajetória permanece em uma faixa vertical em torno de $e = e_1$ no plano de fase, a não linearidade pode ser substituída por uma constante $n'(e_1) = K$. O comportamento das trajetórias em uma faixa em torno de $e = e_1$ será governado por uma

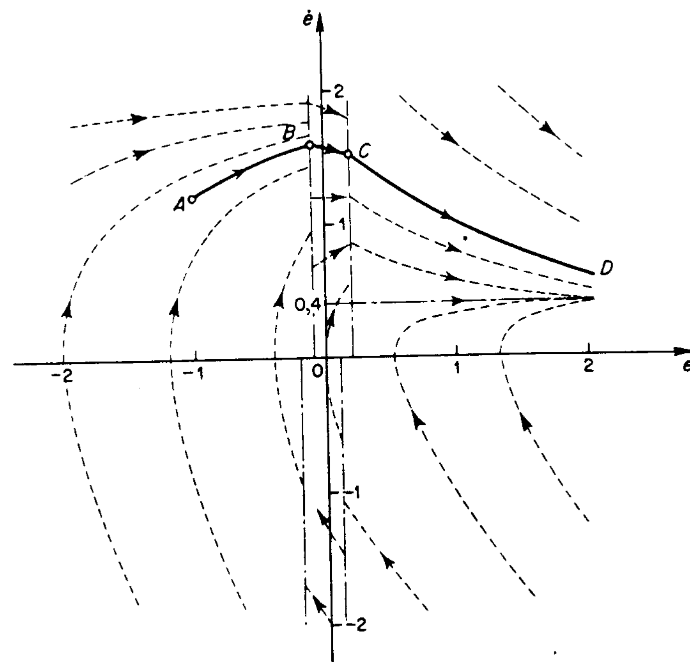


Fig. 12.37 Gráfico de plano de fase do sistema visto na Fig. 12.35(a). [$r(t) = 1{,}2t$].

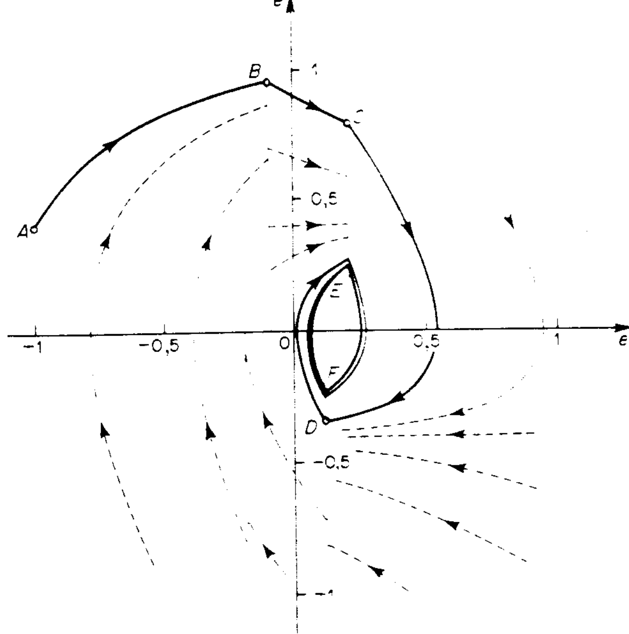


Fig. 12.38 Gráfico de plano de fase do sistema visto na Fig. 12.35(a). [$r(t) = 0.4t$].

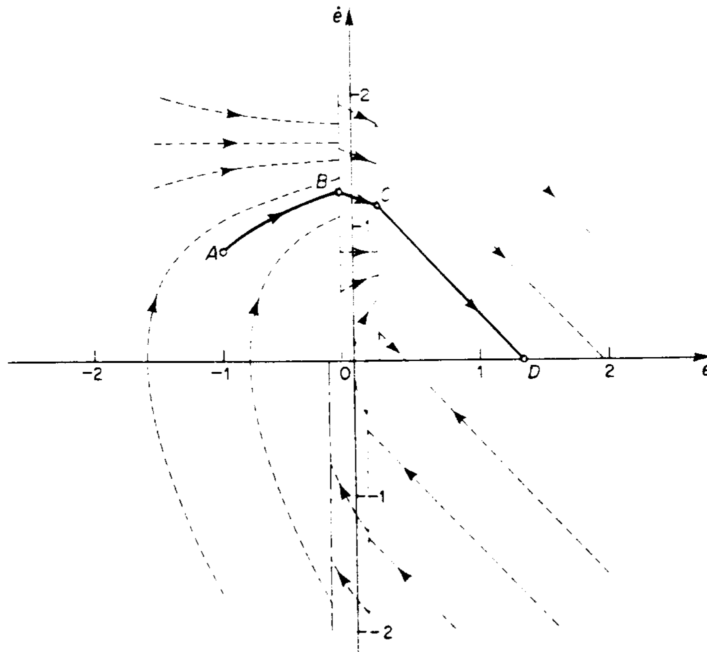


Fig. 12.39 Gráfico de plano de fase do sistema visto na Fig. 12.35(a). [$r(t) = 0.8t$].

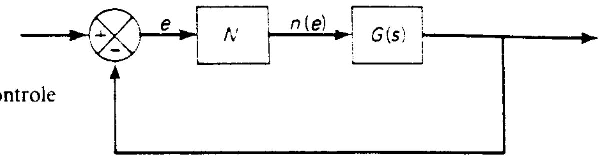


Fig. 12.40 Sistema de controle não linear.

equação diferencial linear invariante no tempo. Isto é, para cada segmento de reta da curva $n(e)$ em função de e corresponde uma região no plano de fase dentro da qual as soluções transitórias obedecem a uma equação diferencial linear. Portanto, o plano de fase inteiro está dividido em várias regiões.

1. Se $n'(e)$ é substituído pelo ganho variável K correspondendo a todos os possíveis valores de $n'(e)$, e se determina que o sistema a malha fechada é estável para todos os valores de K, então é intuitivamente claro que o sistema é estável e que todas as trajetórias irão convergir para um ponto singular único e estável.
2. Se o sistema a malha fechada é instável para alguns valores de K e estável para outros, o sistema pode apresentar um ciclo-limite, ou ciclos-limite. Considere o caso em que todo o plano de fase é dividido em três regiões, como visto na Fig. 12.41. Se as trajetórias na dada região têm um ponto singular dentro desta região, então ele é um ponto singular real. Se o ponto singular está localizado fora da região dada, ele é um ponto singular virtual. Referindo-se à Fig. 12.41, suponha que a região II tem um ponto singular real e que as regiões I e III têm pontos singulares virtuais. Suponha também que o ponto singular real é instável e que os pontos singulares virtuais são estáveis. Pode-se ver que todas as trajetórias irão entrar na região II. Como o ponto singular nesta região é instável, as trajetórias não podem permanecer na região II, e acabarão tendo que deixá-la. Portanto, pode-se ver que as trajetórias não podem terminar em um ponto singular pois o único ponto singular real é instável. Elas não podem tender a infinito pois os pontos singulares virtuais são estáveis. Portanto, o único comportamento possível para as trajetórias é tender a um ciclo-limite estável. O conceito de pontos singulares reais e virtuais é muito fácil de ser tratado, e pode ser estendido facilmente para sistemas de ordem maior que a segunda.

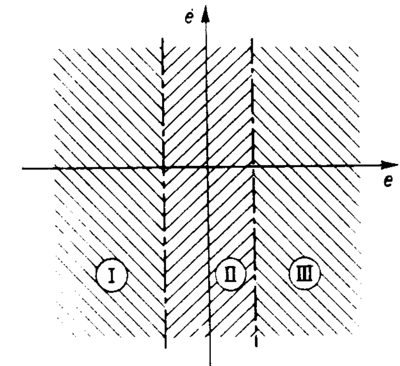


Fig. 12.41 Plano de fase dividido em três regiões.

Um sistema de equações diferenciais ordinárias de ordem n contendo um número arbitrário de não linearidades contínuas e unívocas, onde cada não linearidade pode ser aproximada por um número suficientemente grande de segmentos de reta (em outras palavras, onde os aspectos topológicos da solução de uma equação diferencial são preservados sempre que a não linearidade é substituída por uma aproximação por uma reta adequadamente escolhida), é chamado *monoestável* se a equação diferencial linear em cada ponto no espaço de fase é estável. Em outras palavras, a equação diferencial a que nos referimos deixa de ser monoestável apenas se a equação diferencial linear incremental é instável em alguma região no espaço de fase. O comportamento dinâmico de sistemas monoestáveis é essencialmente similar ao de sistemas lineares. Se a trajetória no espaço de fase permanece dentro de duas superfícies concêntricas esféricas com nenhum ponto singular disponível, então não podemos afirmar que a trajetória tenderá para uma órbita periódica ou ciclo-limite. Entretanto é verdade que a trajetória tenderá para uma trajetória limite quase periódica. Em um sistema de controle, tal trajetória geralmente representa um desempenho indesejável, tanto quanto uma órbita periódica.

12.7 COMENTÁRIOS CONCLUSIVOS

Neste capítulo, apresentamos a análise gráfica associada com sistemas de segunda-ordem. A maior parte da discussão foi dada no plano de fase com as coordenadas x e $\dot{x}$. Outras variáveis podem ser usadas como coordenadas, mas o gráfico de plano de fase se modificará de acordo.

Mostramos que o comportamento limite das trajetórias de um sistema de segunda-ordem quando o tempo t tende a infinito será uma das três possibilidades seguintes:

1. As trajetórias tendem para um ou mais pontos de equilíbrio estáveis.
2. As trajetórias tendem a infinito.
3. As trajetórias tendem a um ciclo-limite ou ciclos-limite.

Mostramos que é possível estabelecer todas as características qualitativas das trajetórias examinando o campo de direções definido pelas equações diferenciais em um número suficientemente grande de pontos no plano de fase. Pode-se fazer com que a precisão de tal gráfico seja tão boa quanto desejada, dentro de certos limites.

Embora a análise gráfica do plano de fase seja limitada a sistemas de segunda-ordem, os conceitos da análise de plano de fase podem ser estendidos para sistemas de ordem superior. Em virtude das dificuldades de trabalhar graficamente em espaço tridimensional e da impossibilidade de visualizar trajetórias em espaços n-dimensionais se $n > 3$, torna-se necessário utilizar outros métodos. Os métodos de análise de espaço de estados, que são indispensáveis na teoria de controle moderno e que são apresentados nos Caps. 14 a 16, são extensões do método do plano de fase e são aplicáveis na análise e síntese de sistemas dinâmicos em espaço n-dimensional.

PROBLEMAS ILUSTRATIVOS E SOLUÇÕES

Problema A.12.1 Considere o pêndulo simples visto na Fig. 12.42. A equação para este sistema é

$$\ddot{\theta} = -\frac{g}{l}\,\text{sen}\,\theta$$

Obtenha a equação para a trajetória. Em seguida construa o gráfico de plano de fase.

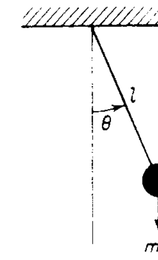


Fig. 12.42 Pêndulo simples.

Solução. Reescrevendo a equação do sistema, obtemos

$$\dot{\theta}\frac{d\dot{\theta}}{d\theta} = -\frac{g}{l}\,\text{sen}\,\theta$$

ou

$$\dot{\theta}\,d\dot{\theta} = -\frac{g}{l}\,\text{sen}\,\theta\,d\theta \qquad (12.61)$$

Integrando ambos os lados da Eq. (12.61), obtemos a equação da trajetória:

$$\frac{1}{2}\dot{\theta}^2 - \frac{g}{l}\cos\theta = k$$

onde k é uma constante. Façamos

$$\frac{d\dot{\theta}}{d\theta} = \alpha, \qquad \sqrt{\frac{g}{l}} = \omega$$

Então a Eq. (12.61) se torna

$$-\omega^2\frac{\text{sen}\,\theta}{\dot{\theta}} = \alpha$$

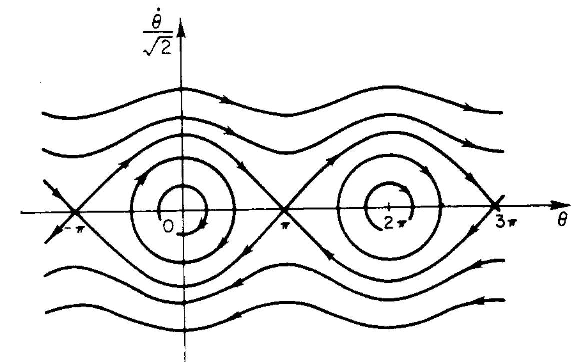


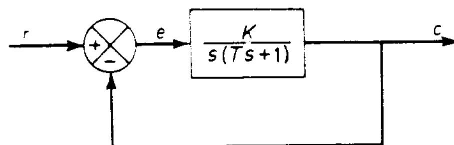
Fig. 12.43 Gráficos de plano de fase para o sistema de pêndulo simples.

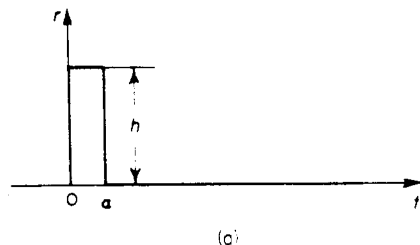
As isóclinas são curvas senoidais. A Fig. 12.43 mostra o gráfico de plano de fase para o sistema.

Problema A.12.2 Considere o sistema visto na Fig. 12.44. Supõe-se que o sistema está inicialmente em repouso. A entrada do sistema é a função pulso vista na Fig. 12.45(a). A altura e a largura do pulso são h e α, respectivamente. Esta entrada em pulso pode ser considerada como a soma de duas entradas em degrau, como visto na Fig. 12.45(b). A entrada em degrau positivo $h\,1(t)$ é aplicada em $t = 0$, e então a entrada em degrau negativo $-h\,1(t-\alpha)$ é aplicada em $t = \alpha$.

Construa a trajetória do sistema quando sujeito a esta entrada em pulso.

Fig. 12.44 Sistema de controle.





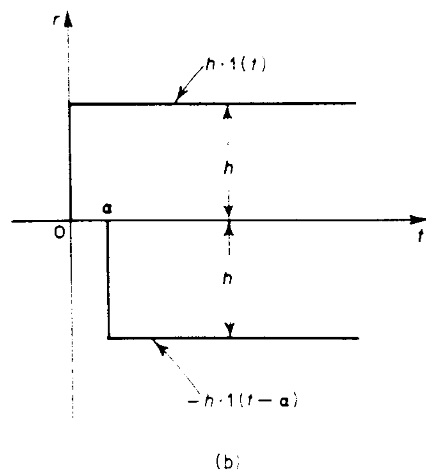


Fig. 12.45 (a) Função pulso; (b) duas funções degrau para se obter a função pulso.

Solução. As equações para o sistema são

$$T\ddot{e} + \dot{e} + Ke = 0 \qquad \text{para } 0 < t < \alpha$$
$$T\ddot{e} + \dot{e} + Ke = -\infty \qquad \text{para } t = \alpha$$
$$T\ddot{e} + \dot{e} + Ke = 0 \qquad \text{para } \alpha < t$$

Suporemos que os pólos da função de transferência de malha-fechada são complexos conjugados e estão no semiplano esquerdo. (Outros casos podem ser tratados de forma semelhante.)

Como é suposto que o sistema está inicialmente em repouso, as condições iniciais para o sinal de erro são $e(0) = h$ e $\dot{e}(0) = 0$. Então, como visto na Fig. 12.46(a), a trajetória no plano $e-\dot{e}$ começa no ponto A e segue o caminho que converge para o foco estável (0, 0) até que $t = \alpha$. Em $t = \alpha$, a trajetória alcança o ponto B. Como a entrada em degrau negativo é aplicada

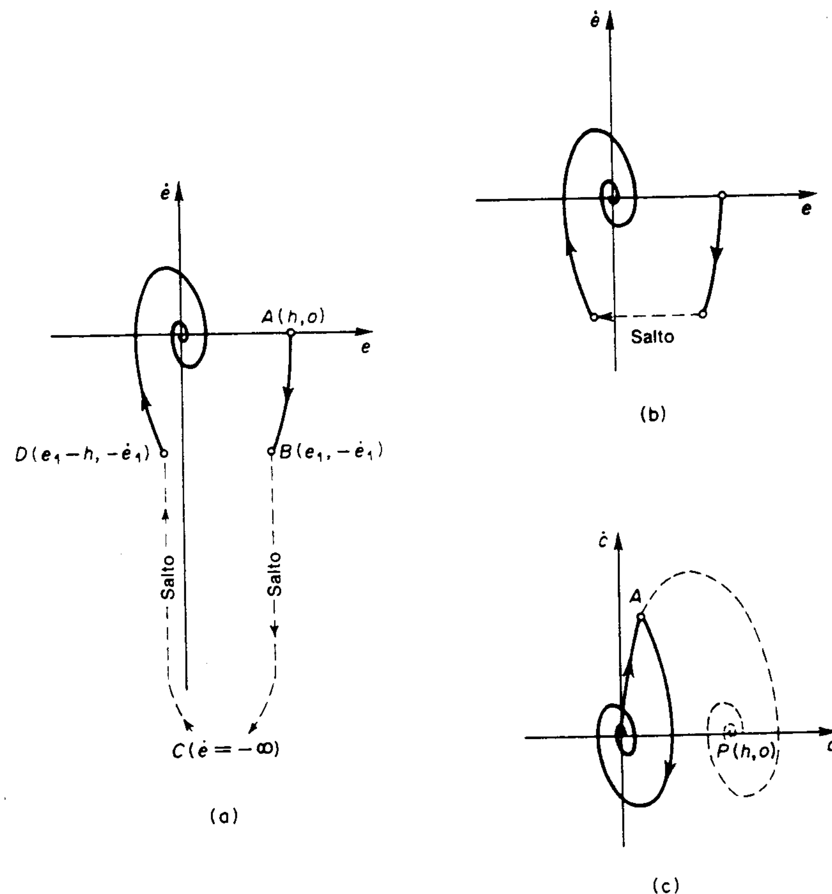


Fig. 12.46 (a) Trajetória no plano e-$\dot{e}$, (b) trajetória indicando um salto, (c) trajetória no plano c-$\dot{c}$.

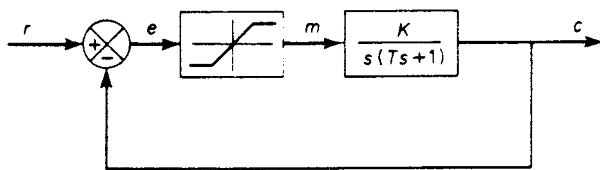
neste instante, a trajetória pula do ponto B ao ponto C, onde $\dot{e} = -\infty$. Do ponto C, a trajetória pula para o ponto D. Estes saltos do ponto B ao ponto C ao ponto D ocorrem no instante $t = \alpha$. Do ponto D, a trajetória converge para a origem, como visto na Fig. 12.46(a). No plano de fase, tais saltos podem ser indicados por uma linha tracejada, como visto na Fig. 12.46(b).

Em termos do sinal de entrada, as equações do sistema se tornam

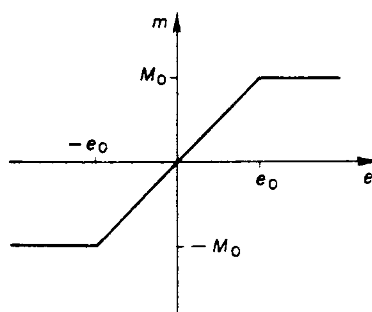
$$T\ddot{c} + \dot{c} + Kc = Kh \qquad \text{para } 0 < t < \alpha$$
$$T\ddot{c} + \dot{c} + Kc = 0 \qquad \text{para } \alpha < t$$

No plano $c-\dot{c}$, a trajetória parte da origem como visto na Fig. 12.46(c). Isto acontece porque as condições iniciais são $c(0) = \dot{c}(0) = 0$. A trajetória tende a convergir para um foco estável P. Em $t = \alpha$, a trajetória alcança o ponto A. Neste momento, o ponto singular do sistema é chaveado do ponto P à origem. Portanto, para $t > \alpha$, a trajetória converge para a origem, um foco estável, como visto na Fig. 12.46(c). A trajetória no plano $c-\dot{c}$ não exibe quaisquer saltos. Como visto da Fig. 12.46(c), a resposta ao pulso da saída do sistema $c(t)$ é oscilatória uma vez que se supõe que os pólos da função de transferência de malha-fechada são complexos conjugados e localizados no semiplano esquerdo. Se $c(t)$ fosse desenhado em função do tempo, o ponto A na Fig. 12.46(c) corresponderia ao primeiro ponto de inflexão da curva de resposta temporal.

Problema A.12.3 A Fig. 12.47(a) mostra um sistema de controle com uma não linearidade tipo saturação. A curva da característica de entrada-saída da não linearidade tipo saturação é vista na Fig. 12.47(b). Supondo que o sistema está inicialmente em repouso, construa as trajetórias no plano de fase quando o sistema é sujeito a uma entrada em degrau $r(t) = R$ e uma entrada em rampa $r(t) = Vt$, $(V > 0)$. Supõe-se que as constantes do sistema são $T = 1$, $K = 4$, $e_0 = 0{,}2$, e $M_0 = 0{,}2$.



(a)



(b)

Fig. 12.47 (a) Sistema de controle com não linearidade tipo saturação. (b) curva da característica entrada-saída da não linearidade tipo saturação.

Solução. Da curva característica da saturação vista na Fig. 12.47(b), obtemos

$$m = e \qquad \text{para } |e| \leq e_0$$
$$m = M_0 \qquad \text{para } e > e_0$$
$$m = -M_0 \qquad \text{para } e < -e_0$$

A equação para o sistema é

$$T\ddot{e} + \dot{e} + Km = T\ddot{r} + \dot{r}$$

Para entrada em degrau: Para a entrada em degrau, $\ddot{r} = \dot{r} = 0$ para $t > 0$. Portanto,

$$T\ddot{e} + \dot{e} + Km = 0$$

Para a operação linear do sistema,

$$T\ddot{e} + \dot{e} + Ke = 0$$

O ponto singular (0, 0) ou é um nó estável ou um foco estável. Para a operação não linear do sistema,

$$T\ddot{e} + \dot{e} + KM_0 = 0 \qquad \text{para } e > e_0$$
$$T\ddot{e} + \dot{e} - KM_0 = 0 \qquad \text{para } e < -e_0$$

Definamos

$$\frac{d\dot{e}}{de} = \alpha$$

e então obtemos

$$\dot{e} = \frac{-\dfrac{KM_0}{T}}{\alpha + \dfrac{1}{T}} \qquad \text{para } e > e_0 \tag{12.62}$$

$$\dot{e} = \frac{\dfrac{KM_0}{T}}{\alpha + \dfrac{1}{T}} \qquad \text{para } e < -e_0 \tag{12.63}$$

Da Eq. (12.62), pode-se ver que para $e > e_0$, todas as trajetórias são assintóticas à reta

$$\dot{e} = -KM_0$$

que corresponde a $\alpha = 0$. Similarmente, da Eq. (12.63), para $e < -e_0$, todas as trajetórias são assintóticas à reta

$$\dot{e} = KM_0$$

A Fig. 12.48 mostra um gráfico de plano de fase para a região $|e| > e_0$. A Fig. 12.49 mostra a trajetória quando o sistema está sujeito a uma entrada em degrau de amplitude 2.

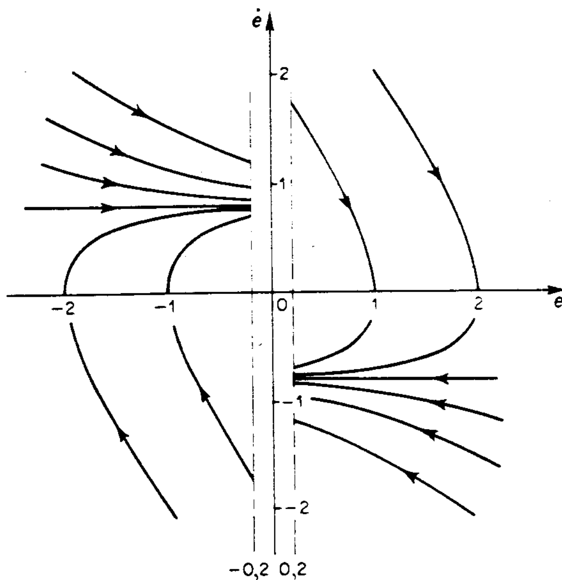


Fig. 12.48 Gráfico de plano de fase para a região $e > e_0$ para o sistema visto na Fig. 12.47(a).

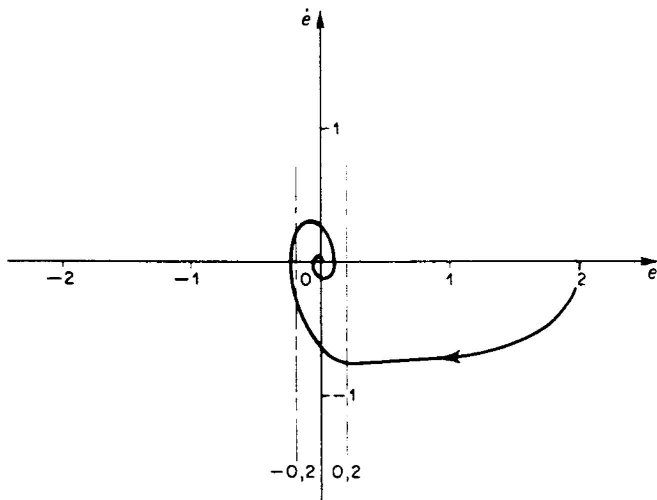


Fig. 12.49 Trajetória correspondendo a uma resposta a degrau do sistema visto na Fig. 12.47(a).

Para entradas em rampa: Para a entrada em rampa, $r = Vt$ e $\dot{r} = V$ = constante. Portanto a equação do sistema se torna

$$T\ddot{e} - \dot{e} - Km = V$$

ou

$$T\ddot{e} - \dot{e} + Ke = V \qquad \text{para } |e| < e_0$$
$$T\ddot{e} - \dot{e} + KM_0 = V \qquad \text{para } e > e_0$$
$$T\ddot{e} - \dot{e} - KM_0 = V \qquad \text{para } e < -e_0$$

O ponto singular para a operação linear está localizado em $(V/K, 0)$. Ele ou é um foco estável ou um nó estável. Para operações não lineares,

$$\dot{e} = \frac{\dfrac{V}{T} - \dfrac{KM_0}{T}}{\alpha + \dfrac{1}{T}} \qquad \text{para } e > e_0$$

$$\dot{e} = \frac{\dfrac{V}{T} + \dfrac{KM_0}{T}}{\alpha + \dfrac{1}{T}} \qquad \text{para } e < -e_0$$

Para $e > e_0$, exceto para o caso especial quando $V = KM_0$, as trajetórias são assintóticas à reta

$$\dot{e} = V - KM_0 \tag{12.64}$$

e, para $e < -e_0$, as trajetórias são assintóticas à reta

$$\dot{e} = V - KM_0 \tag{12.65}$$

Das Eqs. (12.64) e (12.65), pode-se ver que a reta para a qual as trajetórias se aproximam assintoticamente ou está acima do eixo e ou abaixo do eixo e, dependendo de $V > KM_0$ ou $V < KM_0$.

Caso 1 ($V > KM_0$): O gráfico de plano de fase para o caso em que $V = 1.2$ é visto na Fig. 12.50. Se a condição inicial é dada pelo ponto A, então a trajetória segue o caminho $ABCD$. Do ponto B ao ponto C, a trajetória tende a convergir para um foco estável, $(V/K, 0)$; entretanto, a trajetória não pode convergir para este foco estável. Ao invés disto, a trajetória se torna assintótica à reta horizontal $\dot{e} = 0.4$. Em regime estacionário, o erro se torna infinito.

Caso 2 ($V < KM_0$): O gráfico de plano de fase para o caso em que $V = 0.4$ é visto na Fig. 12.51. Se a condição inicial é dada pelo ponto A, então a trajetória converge para um foco estável, $(V/K, 0)$, o ponto singular real.

Caso 3 ($V = KM_0$): Neste caso, para $e > e_0$,

$$T\ddot{e} - \dot{e} = 0$$

ou

$$\dot{e}\left(T\frac{d\dot{e}}{de} - 1\right) = 0$$

Isto significa que para $e > e_0$ as trajetórias ou são retas com inclinação $-1/T$ ou a reta $\dot{e} = 0$. O gráfico de plano de fase para o caso em que $V = 0.8$ é visto na Fig. 12.52. Se a condição inicial é

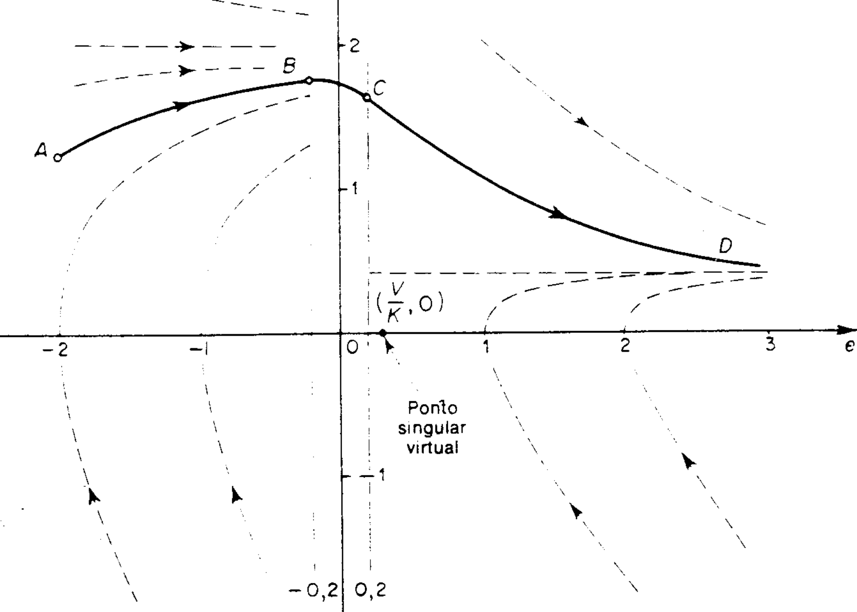


Fig. 12.50 Gráfico de plano de fase do sistema visto na Fig. 12.47(a). [$r(t) = 1,2t$].

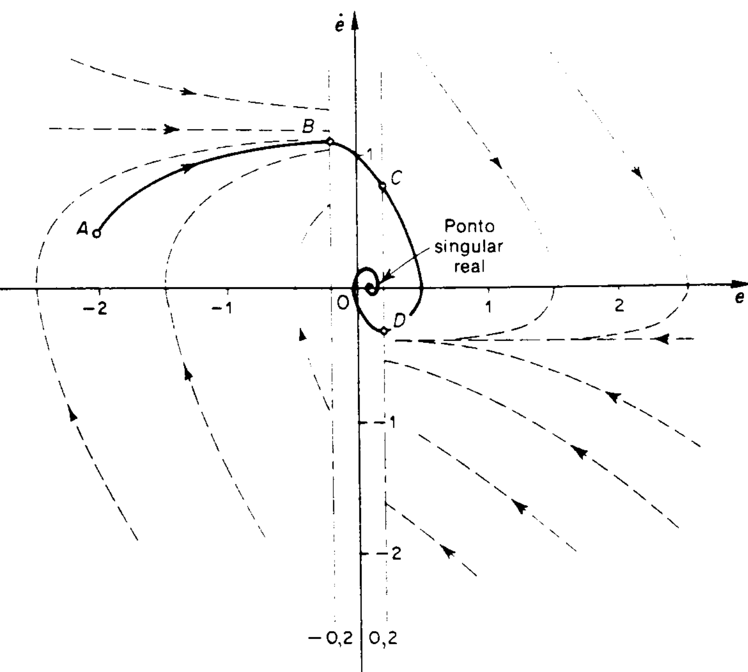


Fig. 12.51 Gráfico de plano de fase do sistema visto na Fig. 12.47(a). [$r(t) = 0,4t$].

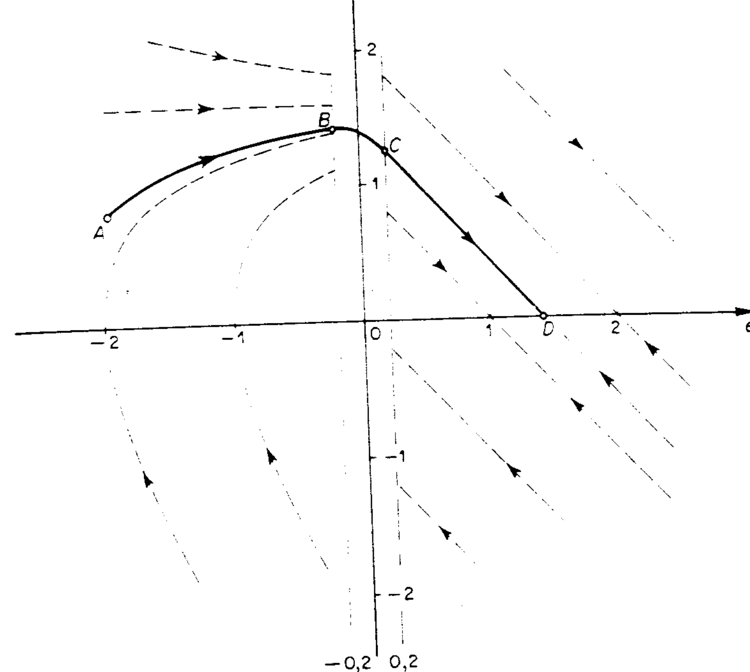


Fig. 12.52 Gráfico de plano de fase do sistema visto na Fig. 12.47(a). [$r(t) = 0,8t$].

dada pelo ponto A, a trajetória segue o caminho $ABCD$. Do ponto B ao ponto C a trajetória tende a convergir para o foco estável em $(V/K, 0)$. Do ponto C a trajetória converge ao ponto D. A resposta termina no ponto D. A amplitude do erro estacionário é $\overline{OD}$.

Problema A.12.4* A Fig. 12.53(a) mostra um diagrama de blocos de um sistema de controle com amortecimento não linear. A curva da característica de entrada-saída do elemento não linear G_N é vista na Fig. 12.53(b). A característica do elemento não linear é tal que o sinal m é nulo se a magnitude do sinal de erro é maior do que e_0. Isto significa que para sinais de erro grandes o sistema tem amortecimento nulo ou uma constante de velocidade infinita. Para sinais de erro menores ($|e| < e_0$), o sistema tem o termo de amortecimento $K_0\dot{c}$. Portanto, a quantidade de sinal tacométrico de realimentação é controlada não linearmente. As características de resposta transitória deste sistema são, em geral, superiores às de sistemas lineares, no sentido de que os primeiros têm características de resposta mais rápida com menor sobre-sinal do que os últimos. (Para um sinal de erro grande, o sistema apresenta resposta mais rápida; para um sinal de erro menor, o sistema é bem amortecido.)

Supondo que o sistema está inicialmente em repouso e que as constantes do sistema são $K = 4$, $K_0 = 1$, e $e_0 = 0,2$, construa a trajetória no plano de fase normalizado quando o sistema é sujeito a uma entrada degrau unitário. Em seguida construa as trajetórias no plano de fase normalizado quando o sistema é sujeito a uma entrada $r(t) = 0,5 - 0,1t$ e uma entrada $r(t) = t$.

*Referência L-8.

Solução. Da Fig. 12.53(a), obtemos

$$\ddot{c} = Kb$$
$$b = e - m\dot{c}$$
$$c = r - e$$

A equação para o sinal de erro pode ser escrita como

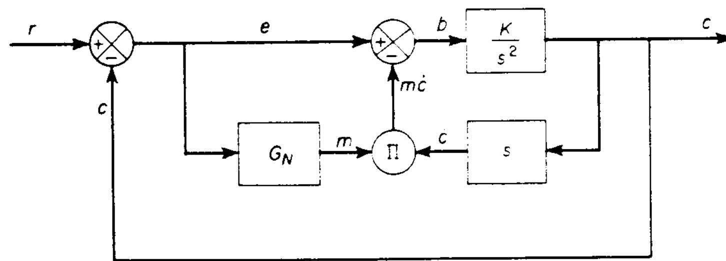
$$\ddot{e} - mK\dot{e} + Ke = \ddot{r} + mK\dot{r} \qquad (12.66)$$

Da Fig. 12.53(b),

$$m = K_0 \qquad \text{para } |e| < e_0$$
$$m = 0 \qquad \text{para } |e| > e_0$$

Para entradas em degrau: Como para a entrada em degrau $\ddot{r} = \dot{r} = 0$ para $t > 0$, a Eq. (12.66) pode ser simplificada para

$$\ddot{e} + KK_0\dot{e} + Ke = 0 \qquad \text{para } |e| < e_0 \qquad (12.67)$$
$$\ddot{e} + Ke = 0 \qquad \text{para } |e| > e_0 \qquad (12.68)$$



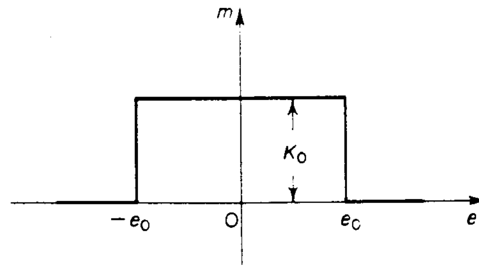


Fig. 12.53 (a) Sistema de controle com amortecimento não linear; (b) curva da característica entrada-saída do elemento não linear G_N.

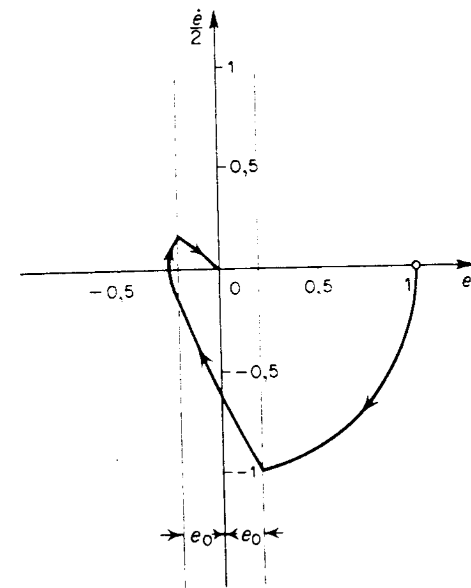


Fig. 12.54 Trajetória correspondendo a uma resposta a degrau unitário do sistema visto na Fig. 12.53(a).

Portanto, a operação do sistema consiste em dois modos lineares. Das Eqs. (12.67) e (12.68), pode-se ver que para ambos os modos de operação linear a origem é o ponto singular. Para $|e| < e_0$, o ponto singular ou é um nó estável ou um foco estável. Para $|e| > e_0$, o ponto singular é um centro.

Da Eq. (12.68), pode-se ver que a trajetória é para $|e| > e_0$ é uma circunferência no plano de fase normalizado [plano $e-(\dot{e}/\sqrt{K})$]. A Fig. 12.54 mostra a trajetória do sinal de erro no plano de fase normalizado quando o sistema é sujeito a uma entrada degrau unitário.

Para entradas em rampa ou rampa mais degrau: Vamos chamar

$$r(t) = R + V(t)$$

Para $t > 0$, a Eq. (12.66) pode ser reescrita como

$$\ddot{e} + mK\dot{e} + Ke = mKV$$

ou

$$\ddot{e} + KK_0\dot{e} + Ke = KK_0V \qquad \text{para } |e| < e_0 \qquad (12.69)$$
$$\ddot{e} + Ke = 0 \qquad \text{para } |e| > e_0 \qquad (12.70)$$

Das Eqs. (12.69) e (12.70), determinamos que o ponto singular para $|e| < e_0$ é $(K_0V, 0)$ e que para $|e| > e_0$ é a origem (0, 0). Como o sistema tem amortecimento positivo para $|e| < e_0$, o ponto singular $(K_0V, 0)$ ou é um nó estável ou um foco estável. Para valores menores de V, ou seja, $V < e_0/K_0$, o ponto singular está no eixo e entre 0 e e_0; para valores maiores de V, ou seja, $V > e_0/K_0$, ele se localiza sobre o eixo e à direita do ponto $e = e_0$.

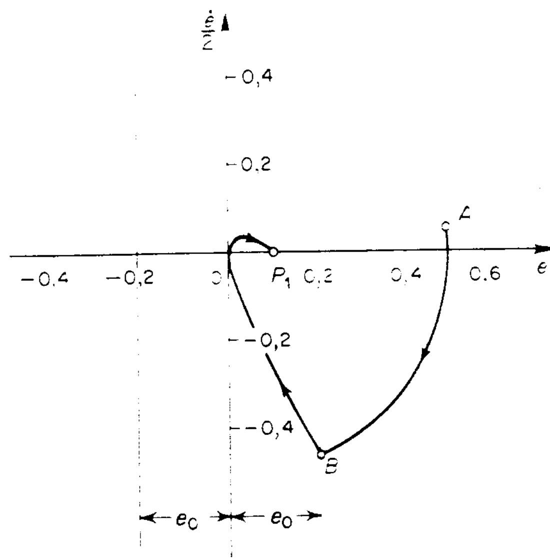


Fig. 12.55 Trajetória correspondendo a uma resposta à rampa (rampa + degrau) do sistema visto na Fig. 12.53(a). [$r(t) = 0.5 + 0.1t$].

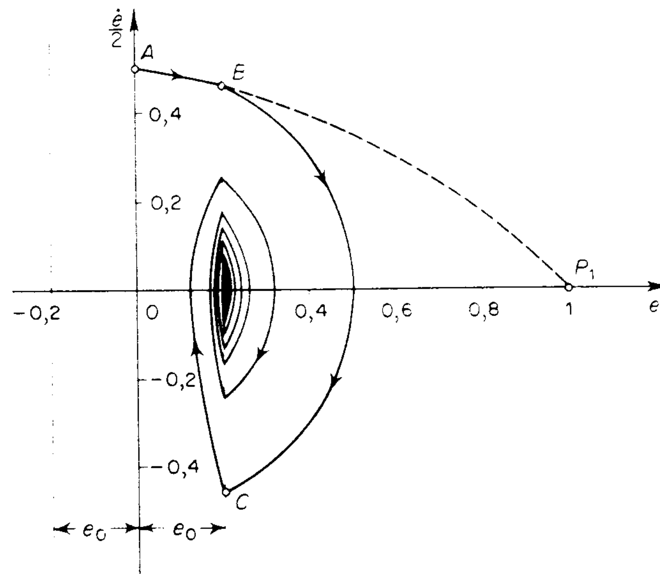


Fig. 12.56 Trajetória correspondendo a uma resposta à rampa do sistema visto na Fig. 12.53(a). [$r(t) = t$].

A Fig. 12.55 mostra a trajetória do sinal de erro no plano de fase normalizado quando $R = 0.5$ e $V = 0.1$. A trajetória parte do ponto A, que é especificado pelas condições iniciais $e(0) = 0.5$, $\dot{e}(0)/2 = 0.05$. O ponto representativo se move ao longo do arco centrado na origem até que a trajetória alcança a reta $e = e_0$ no ponto B, onde a operação do sistema é chaveada para a outra operação linear. Deste ponto em diante, a trajetória converge para o ponto singular P_1, $(K_0V, 0)$. Em regime estacionário, a amplitude do erro é $\overline{OP}_1$.

A Fig. 12.56 mostra a trajetória do sinal de erro no plano de fase normalizado quando $R = 0$ e $V = 1$. A trajetória começa no ponto A, que corresponde às condições iniciais $e(0) = 0$, $\dot{e}(0)/2 = 0.5$, e começa a convergir para o ponto singular P_1, $(1, 0)$. No ponto B ocorre o chaveamento na operação do sistema, e a trajetória começa a convergir para um centro $(0, 0)$. No ponto C, ocorre outro chaveamento na operação. Repetindo o mesmo processo, a trajetória finalmente converge para o ponto $(e_0, 0)$. Nas vizinhanças do ponto $(e_0, 0)$, o sinal de erro exibe pequenas oscilações.

O comportamento de regime estacionário do sistema depende da construção física deste. Se o chaveamento entre duas operações lineares acontece instantaneamente, o sistema apresentará o *chattering* em regime estacionário. Se há um atraso considerável no chaveamento, entretanto, o sistema exibirá um ciclo limite em torno do ponto $(e_0, 0)$.

Da análise que fizemos, pode-se ver que para a entrada em rampa $r(t) = Vt$ o sistema não linear presente tem um erro estacionário de amplitude e_0 ou menor, dependendo da amplitude de V. Isto pode ser considerado como uma vantagem do sistema não linear em discussão sobre seu sistema linear correspondente, porque, para uma entrada em rampa $r(t) = Vt$, o sistema linear cuja equação diferencial é dada pela Eq. (12.69) tem um erro estacionário de K_0V, que é proporcional a V; portanto ele pode ter um valor grande para um valor grande de V; no caso do sistema não linear em discussão, a amplitude do erro estacionário para entradas em rampa é no máximo e_0, que pode ser feito pequeno.

Problema A.12.5 A Fig. 12.57 mostra um servomecanismo de segunda-ordem com atrito de Coulomb. O atrito do Coulomb é uma força de atrito independente da amplitude da

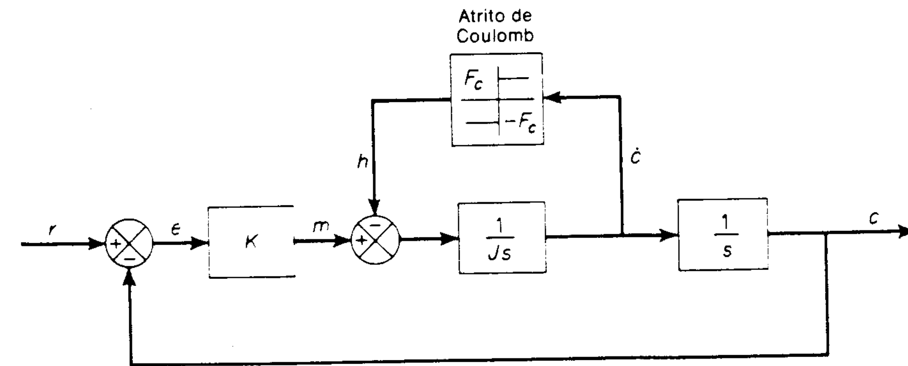


Fig. 12.57 Servomecanismo de segunda-ordem com atrito de Coulomb.

velocidade mas sempre se opondo a ela. Desenhe as trajetórias no plano $e-\dot{e}$ quando o sistema é sujeito às condições iniciais (i) $e(0) = 2.2$, $\dot{e}(0) = 0$; (ii) $e(0) = 3.5$, $\dot{e}(0) = 0$. Suponha que as constantes do sistema são $K = 2$, $J = 1$, e $F_c = 1$.

Solução. As equações que descrevem a dinâmica do sistema são

$$J\ddot{c} = m - h = Ke - h$$
$$c = r - e$$

onde

$$h = F_c \qquad \text{para } \dot{e} > 0$$
$$\;\; = -F_c \qquad \text{para } \dot{e} < 0$$

Em termos do sinal de erro, as equações do sistema podem ser simplificadas fornecendo

$$J\ddot{e} - Ke - h = J\ddot{r} \tag{12.71}$$

Na análise presente, $r = 0$. Portanto a Eq. (12.71) se torna

$$J\ddot{e} - Ke - h = 0$$

ou

$$J\dot{e}\frac{d\dot{e}}{de} - Ke + F_c = 0 \qquad \text{para } \dot{e} > 0 \tag{12.72}$$

$$J\dot{e}\frac{d\dot{e}}{de} - Ke - F_c = 0 \qquad \text{para } \dot{e} < 0 \tag{12.73}$$

Vamos definir

$$\frac{K}{J} = \omega^2$$

Então a Eq. (12.72) pode ser escrita

$$\frac{1}{\omega^2}\dot{e}\,d\dot{e} + \left(e + \frac{F_c}{K}\right) de = 0$$

Integrando ambos os lados desta última equação, obtém-se

$$\left(\frac{\dot{e}}{\omega}\right)^2 - \left(e + \frac{F_c}{K}\right)^2 = A^2 \tag{12.74}$$

onde A é uma constante. Esta é a equação de uma circunferência no plano de fase normalizado [plano $e-(\dot{e}/\omega)$]. O centro da circunferência está em $e = -F_c/K$, $\dot{e}/\omega = 0$, e o raio é igual a A. Como a Eq. (12.74) vale para $\dot{e} > 0$, as trajetórias no semiplano superior no plano de fase normalizado são semicircunferências.

De forma semelhante, a Eq. (12.73) pode ser reescrita como

$$\left(\frac{\dot{e}}{\omega}\right)^2 - \left(e - \frac{F_c}{K}\right)^2 = B^2$$

onde B é uma constante. As trajetórias no semiplano esquerdo do plano de fase normalizado também são semicircunferências.

Pode-se ver que o efeito de atrito de Coulomb na resposta do sistema é deslocar o gráfico de plano de fase do correspondente sistema linear, sem o atrito de Coulomb, da distância F_c/K para a esquerda no semiplano superior e, a mesma distância, para a direita no semiplano inferior.

A Fig. 12.58 mostra duas trajetórias correspondendo às condições iniciais dadas. A trajetória denotada por (i) indica que há um erro estacionário de magnitude 0,2. A trajetória denotada por (ii) termina em $e(0) = -0{,}5$, $\dot{e}(0) = 0$; portanto o sistema exibe um erro estacionário de magnitude 0,5.

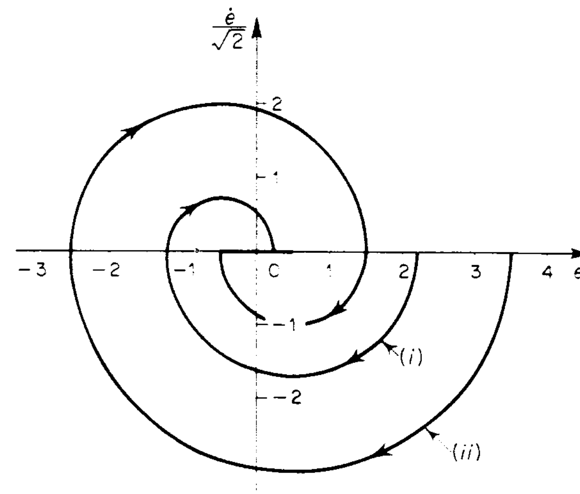


Fig. 12.58 Trajetórias para o servomecanismo visto na Fig. 12.57.

PROBLEMAS

Problema B.12.1 Construa um gráfico de plano de fase do sistema definido por

$$\dot{x}_1 = x_1 - x_2$$
$$\dot{x}_2 = 2x_1 - x_2$$

Problema B.12.2 Construa um gráfico de plano de fase para o seguinte sistema:

$$\ddot{x} + \dot{x} + |x| = 0$$

Problema B.12.3 Construa um gráfico de plano de fase do seguinte sistema:

$$\ddot{\theta} + \dot{\theta} + \operatorname{sen}\theta = 0$$

Problema B.12.4 Determine as localizações e tipos de pontos singulares do sistema não linear descrito por

$$\dot{x}_1 = 0{,}3 - 0{,}1x_1 + x_2 - 0{,}188x_1^2x_2 - 0{,}75x_2^3$$
$$\dot{x}_2 = -0{,}25x_1 - 0{,}1x_2 - 0{,}047x_1^3 + 0{,}188x_1x_2^2$$

Problema B.12.5 A seguinte equação é chamada de equação de Van der Pol:

$$\ddot{x} - (1 - x^2)\dot{x} - x = 0$$

Determine o tipo do ponto singular. Desenhe um gráfico de plano de fase.

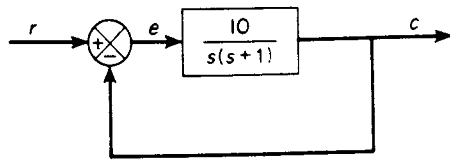


Fig. 12.59 Sistema de controle.

Problema B.12.6 Obtenha a trajetória que representa a resposta do sistema visto na Fig. 12.59 quando este é sujeito à entrada

$$r(t) = R_1\,1(t) + R_2\,1(t-2) + R_3\,1(t-3)$$

onde $1(t - t_i)$ é uma função degrau unitário que ocorre em $t = t_i$. Suponha que o sistema está inicialmente em repouso.

Problema B.12.7 Considere o sistema visto na Fig. 12.60. Construa gráficos de plano de fase para este sistema no plano e-$\dot{e}$ quando $K = 0$ e $K = 1$. Suponha que a entrada $r(t)$ é nula para $t > 0$ e que o sistema está sujeito apenas à condição inicial.

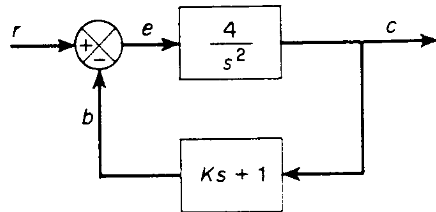


Fig. 12.60 Sistema de controle.

Problema B.12.8 Construa diagramas de plano de fase para o sistema visto na Fig. 12.61 quando $\Delta = 0$ e $\Delta = 0{,}1$. A entrada r é uma função degrau unitário. Use e e $\dot{e}$ como as coordenadas.

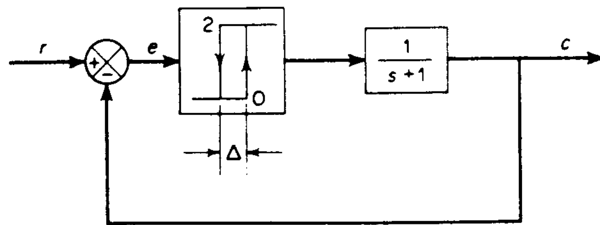


Fig. 12.61 Sistema de controle não linear.

Problema B.12.9 Considere o sistema visto na Fig. 12.62. A entrada r é uma função rampa unitária. Construa uma trajetória típica no plano e-$\dot{e}$.

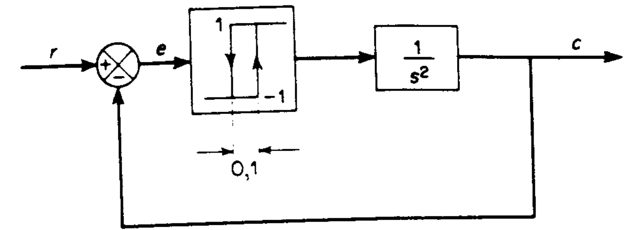


Fig. 12.62 Sistema de controle não linear.

Problema B.12.10 Considere o sistema visto na Fig. 12.63. Suponha que ele está sujeito apenas à condição inicial. Construa uma trajetória típica no plano c-$\dot{c}$.

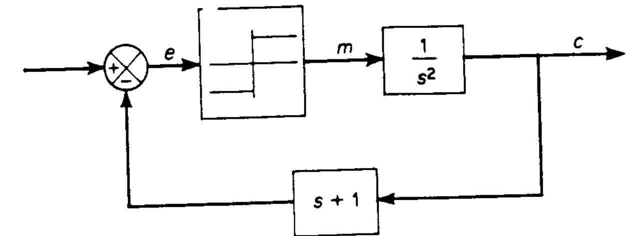


Fig. 12.63 Sistema de controle não linear.

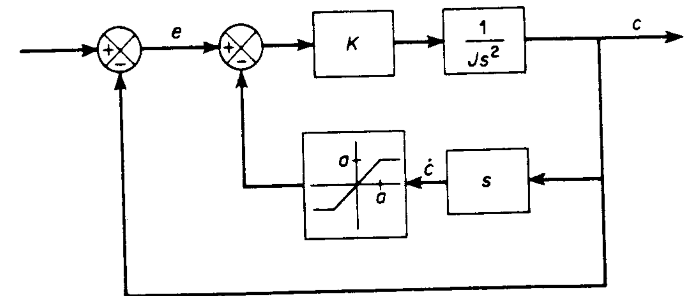


Fig. 12.64 Sistema de controle não linear.

Problema B.12.11 A Fig. 12.64 mostra um sistema de segunda-ordem com um ganho de realimentação não linear. O bloco K indica um ganho de controle proporcional, e o bloco $1/Js^2$ indica uma carga inercial pura. A característica do elemento não linear na malha de realimentação tacométrica é o de uma saturação.

Desenhe trajetórias típicas no plano e-$\dot{e}$ mostrando as respostas para várias condições iniciais. Suponha que $K = 5$, $J = 1$, e $a = 1$.

13

Sistemas de Tempo Discreto e o Método da Transformada z

13.1 INTRODUÇÃO A SISTEMAS DE TEMPO DISCRETO

Sistemas de tempo discreto, ou sistemas a dados amostrados, são sistemas dinâmicos em que uma ou mais variáveis podem mudar apenas em instantes discretos de tempo. Estes instantes, que denotaremos por kT ou t_k ($k = 0, 1, 2, \ldots$), podem especificar o instante em que é feita alguma medida física ou o instante em que é lida a memória de um computador digital etc. O intervalo de tempo entre dois instantes discretos é considerado suficientemente pequeno, de tal forma que os dados para os tempos entre estes instantes discretos podem ser aproximados por interpolação simples.

Sistemas de tempo discreto diferem dos de tempo contínuo, em que os sinais para um sistema de tempo discreto estão na forma amostrada.

Sistemas de tempo discreto ocorrem na prática quando as medidas necessárias para o controle são obtidas de forma intermitente, ou um controlador de larga escala ou computador é multiplexado no tempo entre vários processos de tal forma que um sinal de controle é mandado para cada processo apenas periodicamente ou quando um computador digital é usado para fazer as computações necessárias para controle. Muitos sistemas de controle industrial modernos são sistemas de tempo discreto uma vez que invariavelmente incluem alguns elementos cujas entradas e/ou saídas são discretas no tempo. Às vezes, entretanto, a operação de amostragem, ou discretização, pode ser inteiramente fictícia e introduzida apenas para simplificar a análise de um sistema de controle que na realidade contém apenas elementos contínuos.

Neste capítulo, estudaremos sistemas de tempo discreto onde o sinal que representa o esforço de controle é constante por trechos e muda apenas em instantes discretos no tempo. Como há vários tipos diferentes de operação de amostragem de importância prática, estes serão alistados como segue:

1. Amostragem periódica (convencional): Neste caso, os instantes de amostragem são igualmente espaçados, ou $t_k = kT$ ($k = 0, 1, 2, \ldots$).
2. Amostragem de ordem múltipla: o padrão dos instantes t_k é repetido periodicamente, ou $t_{k+r} - t_k$ = constante para todo k.
3. Amostragem de taxa múltipla: Neste caso, duas amostragens concorrentes ocorrem a $t_k = pT_1$ e qT_2, onde T_1, T_2 são constantes e p, q são inteiros.
4. Amostragem aleatória: Neste caso, os instantes de amostragem são aleatórios, ou t_k é uma variável aleatória.

Neste livro trataremos somente do caso de amostragem periódica.

Quantização. A inclusão de um computador digital em um sistema analógico produz sinais em forma digital (normalmente como números binários) em parte do sistema. O sistema então toma a forma de uma combinação mista digital-analógica. A introdução de um computador digital em um sistema de controle requer o uso de conversores digital-analógico e analógico-digital. A conversão de um sinal analógico para o correspondente sinal digital (número binário) é uma aproximação porque o sinal analógico pode assumir um número infinito de valores, ao passo que a variedade de diferentes números que podem ser formados por um conjunto finito de dígitos é limitada. Este processo de aproximação é chamado de *quantização*.

O processo de quantização (conversão de um sinal em forma analógica para um em forma digital) pode ser ilustrado através da curva característica da Fig. 13.1. A gama de amplitudes de entrada é dividida em um número finito de intervalos disjuntos h_i que não são necessariamente iguais. Todas as amplitudes caindo dentro de cada intervalo são equacionadas a um valor único dentro do intervalo. Este valor único é a aproximação digital para as amplitudes do sinal de entrada analógico. Portanto, se x é a entrada analógica, a saída digital é dada por $y = Q(x)$, onde Q é a função de quantização.

A função $x(t)$ ilustrada na Fig. 13.2(a) é uma função discreta; aquela vista na Fig. 13.2(b) é uma função quantizada, e aquela vista na Fig. 13.2(c) é quantizada tanto em amplitude como no tempo. A operação de sistemas de controle digital envolve quantização tanto em amplitude como em tempo.

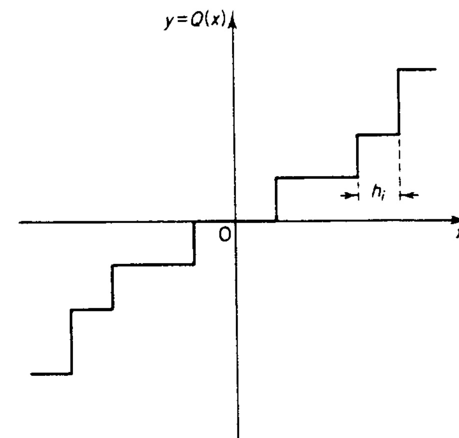


Fig. 13.1 Uma curva mostrando quantização.

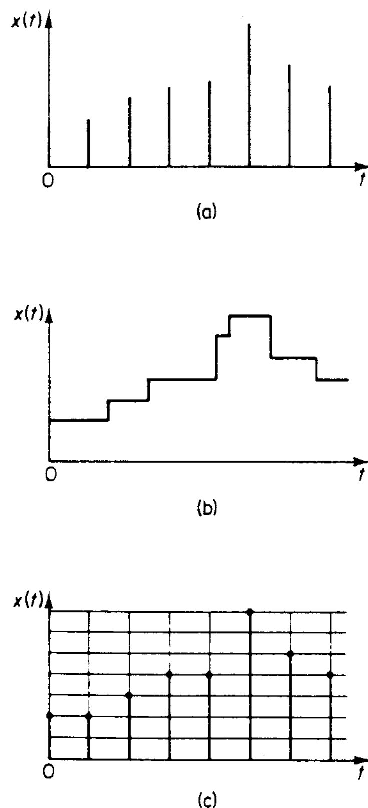


Fig. 13.2 (a) Função de tempo discreto; (b) função quantizada; (c) função quantizada de tempo discreto.

Definições. Apresentaremos em seguida as definições de vários termos.

Transdutor. Um transdutor é um dispositivo que converte um sinal de entrada em um sinal de saída de outra forma. (A saída em geral depende da história passada da entrada.)

Transdutor analógico. Um transdutor analógico é um transdutor em que os sinais de entrada e saída são funções contínuas no tempo. As amplitudes destes sinais podem ter qualquer valor dentro de limitações físicas do sistema.

Transdutor a dados amostrados. Este é um transdutor em que os sinais de entrada e saída ocorrem apenas em instantes discretos de tempo (normalmente periódico), mas as amplitudes do sinal, como no caso do transdutor analógico, são não-quantizadas.

Transdutor digital. Um transdutor digital é aquele em que os sinais de entrada e saída ocorrem apenas em instantes discretos de tempo e as amplitudes dos sinais são quantizadas, *i.e.*, podem assumir apenas certos níveis discretos.

Transdutor analógico-digital. Este é um transdutor em que o sinal de entrada é uma função contínua do tempo e o sinal de saída é um sinal quantizado que pode assumir apenas certos valores discretos.

Transdutor digital-analógico. Um transdutor digital-analógico é aquele em que o sinal de entrada é um sinal quantizado e o sinal de saída é uma função contínua no tempo.

Controladores analógicos e controladores digitais. Ao considerar os tipos de controladores que são usados em sistemas de controle industriais, podemos dividi-los nas seguintes três categorias:

Controladores ou computadores analógicos: Estes representam as variáveis nas equações por grandezas físicas contínuas. Controladores analógicos podem ser projetados para que sirvam satisfatoriamente como controladores não de decisão.

Controladores ou computadores digitais: Estes operam apenas sobre números. A função de decisão é importante em controladores digitais, e atualmente estão sendo usados para a solução de problemas envolvendo a operação ótima global de processos industriais.

Controladores ou computadores analógico-digitais: Estes são freqüentemente chamados de controladores híbridos. São combinações de controladores analógicos e digitais. Alguns dos controladores de alto desempenho são deste tipo.

Vantagens de controladores digitais sobre controladores analógicos. Algumas das vantagens de controladores digitais sobre controladores analógicos podem ser resumidas como segue:

1. Controladores digitais são capazes de realizar complexas computações com precisão constante a grande velocidade. Computadores digitais podem ter quase qualquer grau de precisão nas computações com um aumento relativamente pequeno de custo. Por outro lado, o custo de computadores analógicos aumenta rapidamente quando aumenta a complexidade das computações, caso uma precisão constante deva ser mantida.
2. Controladores digitais são extremamente versáteis. Fazendo um novo programa, podem-se modificar completamente as operações que estão sendo executadas. Esta característica é particularmente importante se o sistema de controle deve receber informação de operação ou instruções de algum centro de computação, onde uma análise econômica e estudos de otimização estão sendo feitos.

Em virtude da inabilidade de técnicas convencionais de tratar problemas de controle complexos, é costumeiro subdividir um processo em unidades menores e então tratar cada uma destas como um problema de controle em separado. Operadores humanos são normalmente usados para coordenar a operação de unidades. Avanços recentes em sistemas de controle por computador causaram mudanças neste uso de controles de processos industriais. Novos desenvolvimentos em computadores de larga escala e de métodos matemáticos fornecem uma base para o uso de toda informação disponível no sistema de controle. No controle convencional, esta parte da malha de controle é feita diretamente por seres humanos.

Controle de sistemas complexos por computador. A tendência atual no controle de sistema de larga escala é de consolidar a multiplicidade de unidades controladas independentemente em processos únicos controlados de forma ótima. Em sistemas de controle de processos industriais, em geral não é prático operar por um tempo muito longo em regime estacionário porque podem ocorrer certas mudanças em especificações de produção, matéria-prima, fatores econômicos, e equipamentos e técnicas de processamento. Portanto, o comportamento transitório de processos industriais deve ser levado em conta. Como há interações entre variáveis do processo, o uso de apenas uma variável de processo para cada agente de controle

não é adequado para um controle realmente completo. Usando controle por computador, é possível levar em conta todas as variáveis do processo juntamente com fatores econômicos, especificações de produção, desempenho de equipamento etc., e portanto obter um controle ótimo de processos industriais.

Note que um sistema capaz de controlar um processo tão completamente quanto possível terá que resolver equações complexas. Quanto mais completo o controle, mais importante se torna conhecer e usar as relações corretas entre variáveis de operação. O sistema deve ser capaz de aceitar instruções de fontes tão variadas como computadores e operadores humanos e também modificar completamente seu subsistema de controle em um tempo curto.

Abordagem de transformada z e de espaço de estado na análise de sistemas de tempo discreto. A análise de sistemas de tempo discreto pode ser feita facilmente em um de dois diferentes métodos. Um é o da transformada z, e o outro é o de espaço de estados.

O método da transformada z tem a mesma relação para sistemas lineares invariantes no tempo de tempo discreto que o método da transformada de Laplace para sistemas lineares de tempo contínuo invariantes no tempo. Este capítulo apresenta apenas o método da transformada z para a análise de sistemas lineares invariantes no tempo de tempo discreto. A abordagem de espaço de estados para a análise de sistemas lineares de tempo discreto é vista nas Seções 14.6 e 14.7.

13.2 A TRANSFORMADA z

Esta seção apresenta o método da transformada z para tratar de funções de tempo discreto. Como dito anteriormente, o papel desempenhado pela transformada z em sistemas de tempo discreto é bastante similar ao da transformação de Laplace em sistemas de tempo contínuo. Como funções de tempo discreto ocorrem quando sinais contínuos são amostrados, iremos primeiro discutir dispositivos amostradores e seguradores.

Dispositivos amostradores e seguradores. O elemento essencial de um sistema de tempo discreto é o amostrador. Em um amostrador convencional, uma chave se fecha para admitir um sinal de entrada a cada T segundos. Na prática, a duração de amostragem é muito curta em comparação com a constante de tempo mais significativa do processo. Um amostrador converte um sinal contínuo em um trem de pulsos ocorrendo nos instantes de amostragem $0, T, 2T, \ldots$, onde T é o período de amostragem. (Entre os instantes de amostragem, o amostrador não transmite nenhuma informação.) Dois sinais, cujos respectivos valores nos instantes de amostragem são iguais, resultarão em um mesmo sinal amostrado.

Um dispositivo segurador converte o sinal amostrado em um sinal contínuo, que reproduz aproximadamente o sinal aplicado a um amostrador. O dispositivo segurador mais simples converte o sinal amostrado em um com amplitude constante entre dois instantes consecutivos de amostragem, como visto na Fig. 13.3. Tal dispositivo é chamado um dispositivo segurador de ordem zero. A função de transferência G_h de um dispositivo segurador de ordem zero é

$$G_h = \frac{1 - e^{-Ts}}{s}$$

Quando o sinal de entrada $x(t)$ é amostrado em instantes discretos, o sinal amostrado é passado pelo dispositivo segurador. Este dispositivo, que é um filtro passa-baixas, alisa o sinal amostrado $x^*(t)$ para produzir o sinal $x_h(t)$, que é constante a partir do último valor amostrado até que a próxima amostra esteja disponível. Isto é,

$$x_h(kT + t) = x(kT) \qquad \text{para } 0 \leq t < T$$

Na análise seguinte, supomos que o dispositivo segurador é de ordem zero. Essencialmente, um dispositivo segurador de ordem zero integra o sinal $x^*(t)$ entre dois instantes consecutivos de amostragem. Notando que a integral de uma função impulso é uma constante, vemos que a entrada de um dispositivo segurador de ordem zero é um trem de funções impulso.

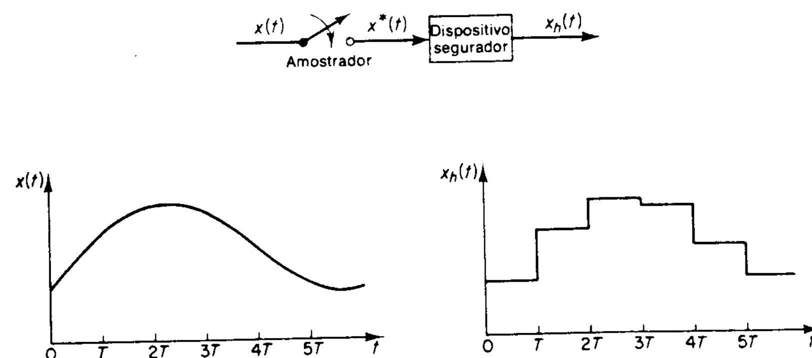


Fig. 13.3 Sinais antes e depois do dispositivo amostrador e segurador.

Considerando a saída do amostrador como um trem de impulsos com certos pesos, podemos relacionar o sinal contínuo $x(t)$ com a saída do amostrador $x^*(t)$ através de

$$x^*(t) = \delta_T(t)\, x(t)$$

onde $\delta_T(t)$ representa um trem de impulsos unitários, como visto na Fig. 13.4(a). A saída do amostrador é igual ao produto da entrada contínua $x(t)$ e o trem de impulsos unitários. Em outras palavras, o amostrador pode ser considerado como um modulador com a entrada $x(t)$ como o sinal modulador e o trem de impulsos unitários como a portadora, como visto na Fig. 13.4(b). Note que $\delta_T(t)$ pode ser escrito como

$$\delta_T(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT)$$

onde $\delta(t - kT)$ é a função impulso unitário ocorrendo em $t = kT$. Se o sinal contínuo $x(t)$ é amostrado de uma forma periódica, o sinal amostrado pode ser representado como

$$x^*(t) = \sum_{k=-\infty}^{\infty} x(t)\, \delta(t - kT) \tag{13.1}$$

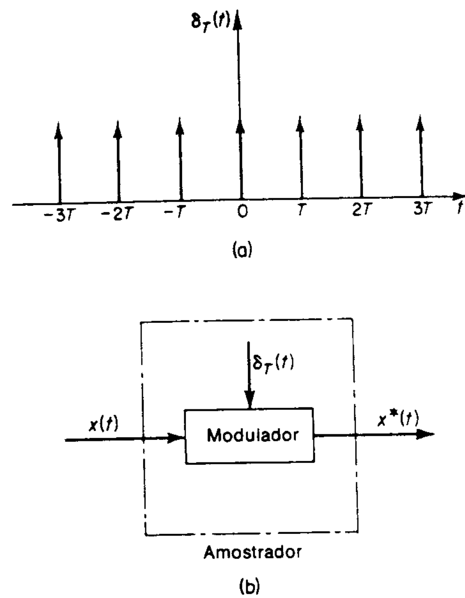


Fig. 13.4 (a) Trem de impulsos unitários; (b) amostrador como um modulador.

ou

$$x^*(t) = \sum_{k=-\infty}^{\infty} x(kT)\,\delta(t-kT) \tag{13.2}$$

A Fig. 13.5 mostra $\delta_T(t)$, $x(t)$, e $x^*(t)$. Como a amplitude de qualquer função impulso é infinita, é conveniente indicar sua magnitude, ou área, pelo comprimento de uma flecha. Na Fig. 13.5, o comprimento de cada flecha no gráfico do sinal amostrado $x^*(t)$ indica a magnitude de cada valor amostrado $x(kT)$.

A maioria das funções temporais que consideraremos neste livro são nulas para $t < 0$. Portanto, a não ser que digamos o contrário, suporemos que este é o caso. Então, para o sinal $x(t)$, as Eqs. (13.1) e (13.2) se tornam, respectivamente,

$$x^*(t) = \sum_{k=0}^{\infty} x(t)\,\delta(t-kT) \tag{13.3}$$

e

$$x^*(t) = \sum_{k=0}^{\infty} x(kT)\,\delta(t-kT) \tag{13.4}$$

Nas análises que se seguem, consideramos a Eq. (13.2) ou (13.4) [alternativamente, a Eq. (13.1) ou (13.3)] como a definição do sinal amostrado $x^*(t)$. Isto é, consideraremos a saída do amostrador como um trem de impulsos, as magnitudes

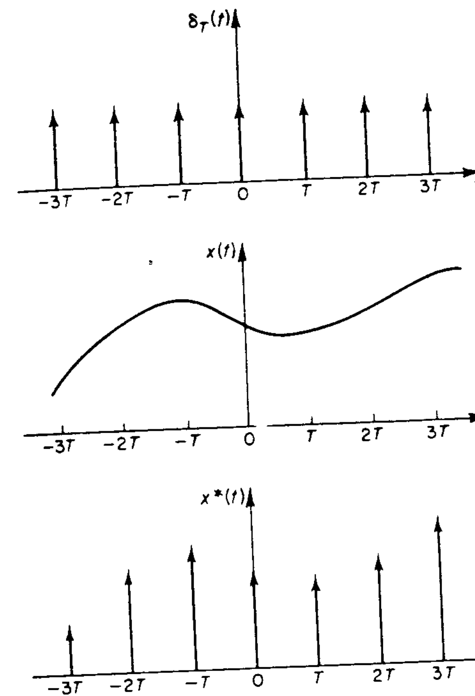


Fig. 13.5 Gráficos de $\delta_T(t)$ em função de t, $x(t)$ em função de t, e $x^*(t)$ *versus* t.

dos quais são iguais aos valores amostrados nos respectivos instantes de amostragem.

Transformada z. Definiremos agora a transformada z. A partir da transformada de Laplace da Eq. (13.4), obtemos

$$X^*(s) = \mathscr{L}[x^*(t)] = \sum_{k=0}^{\infty} x(kT)e^{-kTs}$$

Definamos

$$e^{Ts} = z$$

e escrevamos $X^*(s)$ como $X(z)$. Então,

$$X(z) = X^*(s) = X^*\left(\frac{1}{T}\ln z\right) = \sum_{k=0}^{\infty} x(kT)z^{-k}$$

$X(z)$ é chamada de transformada z de $x^*(t)$, e a notação para a transformada z de $x^*(t)$ é $\mathcal{Z}[x^*(t)]$.

Na transformada z, consideramos apenas os valores do sinal nos instantes de amostragem. Portanto, a transformada z de $x(t)$ e de $x^*(t)$ fornecem o mesmo resultado, ou

$$\mathcal{Z}[x(t)] = \mathcal{Z}[x^*(t)] = X(z) = \sum_{k=0}^{\infty} x(kT)z^{-k} \tag{13.5}$$

Como $X(z)$ depende apenas dos valores de $x(t)$ em $t = kT$ ($k = 0, 1, 2, \ldots$), a transformada inversa z de $X(z)$ nos dá informação sobre $x(t)$ apenas nos instantes de amostragem. (Para a transformada z inversa, referir-se à Seção 13.4.)

Note que a Eq. (13.5) não é a única forma que fornece a transformada z de $x(t)$. Há outras duas expressões para a transformada z. (Veja Problemas A.13.5 a A.13.8.)

Uma tabela de transformadas z comuns é dada na Tabela 13.1. A Tabela 13.2 nos dá propriedades úteis de transformada z.

Exemplo 13.1 Ache a transformada z da função degrau unitário $1(t)$ usando a Eq. (13.5).

Tabela 13.1 Uma tabela de transformadas z

	$X(s)$	$x(t)$ ou $x(k)$	$X(z)$
1	1	$\delta(t)$	1
2	e^{-kTs}	$\delta(t-kT)$	z^{-k}
3	$\frac{1}{s}$	$1(t)$	$\frac{z}{z-1}$
4	$\frac{1}{s^2}$	t	$\frac{Tz}{(z-1)^2}$
5	$\frac{1}{s+a}$	e^{-at}	$\frac{z}{z-e^{-aT}}$
6	$\frac{a}{s(s+a)}$	$1-e^{-at}$	$\frac{(1-e^{-aT})z}{(z-1)(z-e^{-aT})}$
7	$\frac{\omega}{s^2+\omega^2}$	$\operatorname{sen}\omega t$	$\frac{z\operatorname{sen}\omega T}{z^2-2z\cos\omega T+1}$
8	$\frac{s}{s^2+\omega^2}$	$\cos\omega t$	$\frac{z(z-\cos\omega T)}{z^2-2z\cos\omega T+1}$
9	$\frac{1}{(s+a)^2}$	te^{-at}	$\frac{Tze^{-aT}}{(z-e^{-aT})^2}$
10	$\frac{\omega}{(s+a)^2+\omega^2}$	$e^{-at}\operatorname{sen}\omega t$	$\frac{ze^{-aT}\operatorname{sen}\omega T}{z^2-2ze^{-aT}\cos\omega T+e^{-2aT}}$
11	$\frac{s+a}{(s+a)^2+\omega^2}$	$e^{-at}\cos\omega t$	$\frac{z^2-ze^{-aT}\cos\omega T}{z^2-2ze^{-aT}\cos\omega T+e^{-2aT}}$
12	$\frac{2}{s^3}$	t^2	$\frac{T^2z(z+1)}{(z-1)^3}$
13		a^k	$\frac{z}{z-a}$
14		$a^k\cos k\pi$	$\frac{z}{z+a}$

$$\mathcal{Z}[1(t)] = \sum_{k=0}^{\infty} 1(kT)z^{-k}$$
$$= 1 + z^{-1} + z^{-2} + \cdots$$
$$= \frac{z}{z-1}$$

Note que sempre que a série infinita $X(z)$, representando a transformada z de uma função, converge em uma região do plano z, não é necessário especificar os valores de z em que $X(z)$ converge ao usar o método da transformada z para resolver problemas de tempo discreto.

Exemplo 13.2 Obtenha a transformada z do sinal $x(t)$ onde

$$x(t) = 0 \quad \text{para } t < 0$$
$$= e^{-at} \quad \text{para } t \geq 0$$

$$\mathcal{Z}[e^{-at}] = \sum_{k=0}^{\infty} e^{-akT}z^{-k}$$
$$= 1 + e^{-aT}z^{-1} + e^{-2aT}z^{-2} + \cdots$$
$$= \frac{z}{z-e^{-aT}}$$

Tabela 13.2 Propriedades da transformada z

	$x(t)$ ou $x(k)$	$\mathcal{Z}[x(t)]$ ou $\mathcal{Z}[x(k)]$
1	$ax(t)$	$aX(z)$
2	$x_1(t)+x_2(t)$	$X_1(z)+X_2(z)$
3	$x(t+T)$ ou $x(k+1)$	$zX(z)-zx(0)$
4	$x(t+2T)$	$z^2X(z)-z^2x(0)-zx(T)$
5	$x(k+2)$	$z^2X(z)-z^2x(0)-zx(1)$
6	$x(t+kT)$	$z^kX(z)-z^kx(0)-z^{k-1}x(T)-\cdots-zx(kT-T)$
7	$x(k+m)$	$z^mX(z)-z^mx(0)-z^{m-1}x(1)-\cdots-zx(m-1)$
8	$tx(t)$	$-Tz\frac{d}{dz}[X(z)]$
9	$kx(k)$	$-z\frac{d}{dz}[X(z)]$
10	$e^{-at}x(t)$	$X(ze^{aT})$
11	$e^{-ak}x(k)$	$X(ze^{a})$
12	$a^kx(k)$	$X\left(\frac{z}{a}\right)$

Tabela 13.2 (Cont.) Propriedades da transformada z

13	$ka^k x(k)$	$-z\frac{d}{dz}\left[X\left(\frac{z}{a}\right)\right]$
14	$x(0)$	$\lim_{z\to\infty} X(z)$ se o limite existe
15	$x(\infty)$	$\lim_{z\to 1}[(z-1)X(z)]$ se $\frac{z-1}{z}X(z)$ é analítica sobre e fora do círculo unitário
16	$\sum_{k=0}^{\infty} x(k)$	$X(1)$
17	$\sum_{k=0}^{n} x(kT)y(nT-kT)$	$X(z)Y(z)$

Exemplo 13.3 Obtenha a transformada z de $x(t)$ onde

$$x(t) = 0 \qquad \text{para } t < 0$$
$$= \text{sen}\,\omega t \qquad \text{para } t \geq 0$$

Como

$$\mathcal{Z}[e^{-at}] = \frac{z}{z - e^{-aT}}$$

obtemos

$$\begin{aligned}\mathcal{Z}[\text{sen}\,\omega t] &= \mathcal{Z}\left[\frac{e^{j\omega t} - e^{-j\omega t}}{2j}\right] \\ &= \frac{1}{2j}\left(\frac{z}{z - e^{j\omega T}} - \frac{z}{z - e^{-j\omega T}}\right) \\ &= \frac{1}{2j}\,\frac{z(e^{j\omega T} - e^{-j\omega T})}{z^2 - z(e^{j\omega T} + e^{-j\omega T}) + 1} \\ &= \frac{z\,\text{sen}\,\omega T}{z^2 - 2z\cos\omega T + 1}\end{aligned}$$

Exemplo 13.4 Obtenha a transformada z de

$$X(s) = \mathcal{L}[x(t)] = \frac{1}{s(s+1)}$$

Sempre que uma função de s é dada, a correspondente transformada z pode ser obtida inicialmente expandindo-se a função dada em s em frações parciais e então combinando a transformada z de cada termo de fração parcial. Vamos expandir $X(s)$ em frações parciais.

$$X(s) = \frac{1}{s(s+1)} = \frac{1}{s} - \frac{1}{s+1}$$

A transformada z, correspondendo a $1/s$ [ou $1(t)$], é $z/(z-1)$, e a correspondente a $1/(s+1)$ [ou e^{-t}] é $z/(z - e^{-T})$. Portanto,

$$\begin{aligned}\mathcal{Z}[x(t)] = X(z) &= \frac{z}{z-1} - \frac{z}{z - e^{-T}} \\ &= \frac{z(1 - e^{-T})}{(z-1)(z-e^{-T})}\end{aligned}$$

13.3 RESOLUÇÃO DE EQUAÇÕES DE DIFERENÇA USANDO O MÉTODO DA TRANSFORMADA z

A resolução de equações de diferença pelo método da transformada z é muito útil, tal como a solução de equações diferenciais por transformadas de Laplace. Essencialmente, usando o método da transformada z, podemos transformar equações de diferenças em equações algébricas em z. A seguir usaremos a notação simplificada $x(k)$ para denotar $x(kT)$.

Transformada z de $x(k+1)$. A transformada z de $x(k+1)$ é dada por

$$\mathcal{Z}[x(k+1)] = zX(z) - zx(0) \tag{13.6}$$

onde $X(z) = \mathcal{Z}[x(k)]$. Isto pode ser demonstrado como segue:

$$\begin{aligned}\mathcal{Z}[x(k+1)] &= \sum_{k=0}^{\infty} x(k+1)z^{-k} \\ &= \sum_{k=1}^{\infty} x(k)z^{-k+1} \\ &= z\left[\sum_{k=0}^{\infty} x(k)z^{-k} - x(0)\right] \\ &= zX(z) - zx(0)\end{aligned}$$

Note que se $x(0) = 0$, então

$$\mathcal{Z}[x(k+1)] = z\mathcal{Z}[x(k)] \qquad \text{se } x(0) = 0$$

Portanto, se $x(0) = 0$, então a multiplicação da transformada z de uma função $x(k)$ por z corresponde a um deslocamento para diante no tempo de um período.

A Eq. (13.6) pode ser facilmente modificada para se obter a seguinte relação:

$$\begin{aligned}\mathcal{Z}[x(k+2)] &= z\mathcal{Z}[x(k+1)] - zx(1) \\ &= z^2X(z) - z^2x(0) - zx(1)\end{aligned}$$

De forma semelhante,

$$\mathcal{Z}[x(k+m)] = z^mX(z) - z^mx(0) - z^{m-1}x(1) - z^{m-2}x(2) - \cdots - zx(m-1)$$

onde m é um inteiro positivo. Note que quando a equação de diferenças é transformada em uma equação algébrica em z pelo método da transformada z, os dados iniciais são automaticamente incluídos na representação algébrica.

Exemplo 13.5 Resolver a seguinte equação de diferenças usando o método da transformada z:

$$x(k+2) + 3x(k+1) + 2x(k) = 0, \qquad x(0) = 0, \qquad x(1) = 1$$

Tomando a transformada z em ambos os lados desta equação de diferenças, obtemos

$$z^2X(z) - z^2x(0) - zx(1) + 3zX(z) - 3zx(0) + 2X(z) = 0$$

Substituindo os dados iniciais e simplificando, temos

$$X(z) = \frac{z}{z^2 + 3z + 2}$$
$$= \frac{z}{(z+1)(z+2)}$$
$$= \frac{z}{z+1} - \frac{z}{z+2}$$

Notando que

$$\mathcal{Z}[a^k] = \frac{z}{z-a}$$

temos

$$\mathcal{Z}[(-1)^k] = \frac{z}{z+1}, \qquad \mathcal{Z}[(-2)^k] = \frac{z}{z+2}$$

Portanto

$$x(k) = (-1)^k - (-2)^k \qquad (k = 0, 1, 2, \ldots)$$

Exemplo 13.6 Determine a resposta $x(k)$ do seguinte sistema:

$$x(k+2) - 3x(k+1) + 2x(k) = u(k) \tag{13.7}$$

onde

$$x(k) = 0 \quad \text{para } k \le 0$$
$$u(0) = 1$$
$$u(k) = 0 \quad \text{para } k < 0, k > 0$$

Substituindo $k = -1$ na Eq. (13.7), obtemos

$$x(1) = 0$$

Tomando a transformada z da Eq. (13.7) com dados iniciais $x(0) = x(1) = 0$, obtemos

$$(z^2 - 3z + 2)X(z) = U(z)$$

Note que a transformada z da função de excitação $u(k)$ é

$$U(z) = \sum_{k=0}^{\infty} u(k)z^{-k} = 1$$

Portanto

$$X(z) = \frac{1}{z^2 - 3z + 2} = \frac{-1}{z-1} + \frac{1}{z-2}$$

Usando a relação

$$\mathcal{Z}[x(k+1)] = zX(z) - zx(0)$$

e notando que $x(0) = 0$, obtemos

$$\mathcal{Z}[x(k+1)] = zX(z)$$
$$= -\frac{z}{z-1} + \frac{z}{z-2}$$

Como

$$\mathcal{Z}[1^k] = \frac{z}{z-1}, \qquad \mathcal{Z}[2^k] = \frac{z}{z-2}$$

obtemos

$$x(k+1) = -1 + 2^k \qquad (k = 0, 1, 2, \ldots)$$

ou

$$x(k) = -1 + 2^{k-1} \qquad (k = 1, 2, 3, \ldots)$$

Teorema do valor inicial. Se $x(t)$ tem a transformada $X(z)$ e $\lim_{z\to\infty} X(z)$ existe, então o valor inicial $x(0)$ de $x(t)$ ou $x(k)$ é dado por

$$x(0) = \lim_{z\to\infty} X(z) \tag{13.8}$$

Para provar isto, note que

$$X(z) = \sum_{k=0}^{\infty} x(k)z^{-k}$$
$$= x(0) + x(1)z^{-1} + x(2)z^{-2} + \cdots$$

Fazendo $z\to\infty$, obtemos a Eq. (13.8).

Teorema do valor final. Se $x(t)$ tem a transformada z $X(z)$, e se $X(z)$ não tem pólos duplos ou de ordem mais alta na circunferência unitária centrada na origem do plano z e nenhum pólo fora do círculo unitário [esta é a condição para a estabilidade de $X(z)$, ou a condição de que $x(k)$ $(k = 0, 1, 2, \ldots)$ permaneça finita (veja a Seção 13.6)], então o valor final de $x(t)$ ou $x(k)$ é dado por

$$\lim_{k\to\infty} x(k) = \lim_{t\to\infty} x(t) = \lim_{z\to 1} [(z-1)X(z)] \tag{13.9}$$

Para provar isto, note que

$$\mathcal{Z}[x(k)] = X(z) = \sum_{k=0}^{\infty} x(k)z^{-k}$$

$$\mathcal{Z}[x(k+1)] = zX(z) - zx(0) = \sum_{k=0}^{\infty} x(k+1)z^{-k}$$

Portanto,

$$zX(z) - zx(0) - X(z) = (z-1)X(z) - zx(0)$$
$$= \sum_{k=0}^{\infty} x(k+1)z^{-k} - \sum_{k=0}^{\infty} x(k)z^{-k}$$

do qual obtemos

$$(z-1)X(z) = zx(0) + \sum_{k=0}^{\infty} [x(k+1) - x(k)]z^{-k}$$

Em virtude da suposta condição de estabilidade, obtemos, quando $z \to 1$,

$$\lim_{z\to 1} [(z-1)X(z)] = x(0) + x(\infty) - x(0) = x(\infty)$$

que é a Eq. (13.9).

13.4 A TRANSFORMADA z INVERSA

Dado $X(z)$, há três métodos para se obter a transformada z inversa, $x(kT)$ ou $x(k)$. Os três métodos são baseados na expansão por série infinita de potências, na expansão por frações parciais, e a integral de inversão. Ao obter a transformada z inversa, supomos como sempre que a série temporal $x(kT)$ ou $x(k)$ é zero para $k < 0$.

Obtenção da transformada z inversa expandindo $X(z)$ em uma série infinita de potências. Se $X(z)$ é expandido em uma série de potências convergentes, como visto abaixo,

$$X(z) = \sum_{k=0}^{\infty} x(kT)z^{-k}$$
$$= x(0) + x(T)z^{-1} + x(2T)z^{-2} + \cdots + x(kT)z^{-k} + \cdots$$

então os valores de $x(kT)$ podem ser determinados por inspeção.

Se $X(z)$ é dada na forma de uma função racional, a expansão em uma série de potências infinita pode ser feita simplesmente dividindo-se o numerador pelo denominador. Se a série resultante é convergente, os coeficientes de z^{-k} na série são os valores $x(kT)$ da seqüência temporal. Ao se obter os coeficientes por divisão longa, tanto o numerador como o denominador de $X(z)$ devem ser escritos como potências ascendentes de z^{-1}.

Embora o presente método forneça os valores de $x(0)$, $x(T)$, $x(2T)$,... em uma forma seqüencial, é normalmente difícil obter uma expressão para o termo geral a partir de um conjunto de valores de $x(kT)$.

Exemplo 13.7 Determine $x(kT)$ para $k = 0, 1, 2, 3, 4$ quando $X(z)$ é dado por

$$X(z) = \frac{10z}{(z-1)(z-2)}$$

$X(z)$ pode ser escrito como

$$X(z) = \frac{10z^{-1}}{1 - 3z^{-1} + 2z^{-2}}$$

Fazendo uma divisão longa,

$$X(z) = 10z^{-1} + 30z^{-2} + 70z^{-3} + 150z^{-4} + \cdots$$

Esta série infinita converge. Portanto, obtemos por inspeção

$$x(0) = 0$$
$$x(T) = 10$$
$$x(2T) = 30$$
$$x(3T) = 70$$
$$x(4T) = 150$$

Obtenção da transformada z inversa expandindo $X(z)$ em frações parciais. Um método alternativo para a obtenção de $x(kT)$ é baseado na expansão em frações parciais de $X(z)/z$ e a identificação de cada um dos termos usando-se uma tabela de transformadas z. (Tabelas de transformadas z complicadas podem não ser disponíveis. Portanto, os polinômios em z devem ser expandidos em frações parciais antes que as transformadas inversas possam ser obtidas.) Note que a razão pela qual expandimos $X(z)/z$ em frações parciais é que as funções de z que aparecem em tabelas normalmente têm o fator z nos seus numeradores.

Considere $X(z)$ dado por

$$X(z) = \frac{b_0 z^m + b_1 z^{m-1} + \cdots + b_{m-1}z + b_m}{a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n} \qquad (m \leq n)$$

Inicialmente fatoramos o polinômio do denominador de $X(z)$ e determinamos os pólos de $X(z)$. Então expandimos $X(z)/z$ em frações parciais de tal forma que cada um dos termos pode ser facilmente reconhecido em uma tabela de transformadas z. A transformada inversa de $X(z)$ é obtida como a soma das transformadas z inversas das frações parciais.

Exemplo 13.8 Determine $x(kT)$ se $X(z)$ é dada por

$$X(z) = \frac{10z}{(z-1)(z-2)}$$

Inicialmente expandimos $X(z)/z$ em frações parciais como segue:

$$\frac{X(z)}{z} = \frac{10}{(z-1)(z-2)} = \frac{-10}{z-1} + \frac{10}{z-2}$$

Portanto,

$$X(z) = -\frac{10z}{z-1} + \frac{10z}{z-2}$$

Da Tabela 13.1, obtemos

$$\mathcal{Z}^{-1}\left[\frac{z}{z-1}\right] = 1, \qquad \mathcal{Z}^{-1}\left[\frac{z}{z-2}\right] = 2^k$$

Portanto

$$x(kT) = 10(-1 + 2^k) \qquad (k = 0, 1, 2, \ldots)$$

ou

$$\begin{aligned} x(0) &= 0 \\ x(T) &= 10 \\ x(2T) &= 30 \\ x(3T) &= 70 \\ x(4T) &= 150 \\ &\cdots \end{aligned}$$

Claramente, o resultado confere com o $x(kT)$ obtido no Exemplo 13.7.

Obtenção da transformada z inversa por integral de inversão. O terceiro método de determinar a transformada inversa é usando a integral de inversão. Da Eq. (13.5), temos

$$\begin{aligned} X(z) &= \sum_{k=0}^{\infty} x(kT)z^{-k} \\ &= x(0) + x(T)z^{-1} + x(2T)z^{-2} + \cdots + x(kT)z^{-k} + \cdots \end{aligned}$$

Multiplicando ambos os lados desta última equação por z^{k-1}, obtemos

$$X(z)z^{k-1} = x(0)z^{k-1} + x(T)z^{k-2} + x(2T)z^{k-3} + \cdots + x(kT)z^{-1} + \cdots \tag{13.10}$$

Note que devido a

$$z = e^{Ts}$$

se $s = \sigma + j\omega$ é substituído nesta última equação, obtemos $z = e^{T(\sigma+j\omega)}$ ou

$$|z| = e^{T\sigma}, \qquad \angle z = \omega T$$

Se os pólos de $\mathcal{L}[x]$ estão para a esquerda da reta $s = \sigma_1$ no plano s, os pólos de $\mathcal{Z}[x]$ estarão no interior da circunferência cujo centro é a origem e o raio é igual a $e^{T\sigma_1}$ no plano z.

Suponha que integramos ambos os lados da Eq. (13.10) ao longo desta circunferência na direção anti-horária:

$$\oint X(z)z^{k-1}\,dz = \oint x(0)z^{k-1}\,dz + \oint x(T)z^{k-2}\,dz + \cdots + \oint x(kT)z^{-1}\,dz + \cdots$$

Aplicando o teorema de Cauchy, vemos que todos os termos do lado direito desta última equação são nulos exceto um termo:

$$\oint x(kT)z^{-1}\,dz$$

Portanto

$$\oint X(z)z^{k-1}\,dz = \oint x(kT)z^{-1}\,dz$$

do que obtemos

$$x(kT) = \frac{1}{2\pi j}\oint X(z)z^{k-1}\,dz \tag{13.11}$$

A Eq. (13.11) é a integral de inversão para a transformada z. A Eq. (13.11) equivale a dizer que

$$x(kT) = \sum [\text{resíduos de } X(z)z^{k-1} \text{ nos pólos de } X(z)] \tag{13.12}$$

Exemplo 13.9 Obtenha $x(kT)$ usando a integral de inversão quando $X(z)$ é dado por

$$X(z) = \frac{10z}{(z-1)(z-2)}$$

Das Eqs. (13.11) e (13.12),

$$\begin{aligned} x(kT) &= \frac{1}{2\pi j}\oint \left[\frac{10z}{(z-1)(z-2)}z^{k-1}\right]dz \\ &= \frac{1}{2\pi j}\oint \left[-\frac{10z^k}{z-1} + \frac{10z^k}{z-2}\right]dz \\ &= \left(\text{resíduo de } -\frac{10z^k}{z-1} \text{ no pólo } z = 1\right) + \left(\text{resíduo de } \frac{10z^k}{z-2} \text{ no pólo } z = 2\right) \\ &= 10(-1 + 2^k) \qquad (k = 0, 1, 2, \ldots) \end{aligned}$$

13.5 FUNÇÕES DE TRANSFERÊNCIA PULSADA

Esta seção apresenta o material básico necessário para analisar sistemas de tempo discreto, ou sistemas a dados amostrados usando o método da transfor-

mada z.

Ao analisar sistemas de tempo discreto por este método, é importante notar que embora os valores da resposta do sistema em instantes de amostragem sejam corretos, a resposta do sistema obtida pelo método da transformada z pode não mostrar o comportamento correto de resposta temporal do sistema real, a não ser que a função de transferência $G(s)$ da parte contínua do sistema tenha pelo menos dois pólos a mais do que zeros, de tal forma que $\lim_{s\to\infty} sG(s) = 0$.

Teorema da amostragem. O teorema da amostragem de Shannon apresentado aqui é importante no projeto de sistemas de tempo discreto pois dá a mínima freqüência de amostragem necessária para reconstruir o sinal original a partir de um sinal amostrado.

Suporemos que um sinal contínuo $x(t)$ tem o espectro de freqüência como mostrado na Fig. 13.6. Este sinal $x(t)$ não contém quaisquer componentes de freqüência acima de ω_1 rad/s.

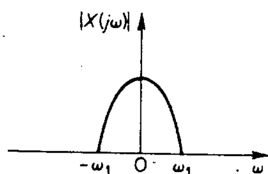


Fig. 13.6 Um espectro de freqüência.

Teorema de amostragem. Se $\omega_s = 2\pi/T$, onde T é o período de amostragem, é maior do que $2\omega_1$, ou

$$\omega_s > 2\omega_1$$

onde $2\omega_1$ corresponde ao espectro de freqüência do sinal contínuo $x(t)$, então o sinal $x(t)$ pode ser completamente reconstruído a partir do sinal amostrado $x^*(t)$.

Demonstraremos isto quando $\lim_{s\to\infty} sX(s) = x(0+) = 0$. Vamos definir

$$X(s) = \mathscr{L}[x(t)]$$

Referindo-se ao Problema A.13.5, a transformada de Laplace do sinal amostrado $x^*(t)$ é dada por

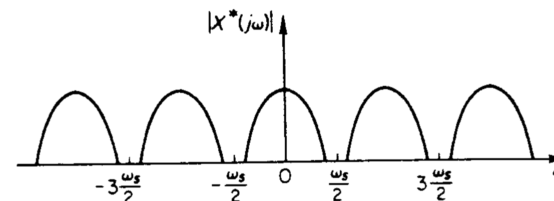
$$X^*(s) = \frac{1}{T}\sum_{k=-\infty}^{\infty} X(s + j\omega_s k)$$

Substituindo $s = j\omega$ nesta última equação, obtemos o espectro de freqüência para $X^*(s)$ como segue:

$$|X^*(j\omega)| = \frac{1}{T}\left|\sum_{k=-\infty}^{\infty} X[j(\omega + \omega_s k)]\right|$$

$$= \cdots + \frac{1}{T}|X[j(\omega - \omega_s)]| + \frac{1}{T}|X(j\omega)|$$

$$+ \frac{1}{T}|X[j(\omega + \omega_s)]| + \cdots$$



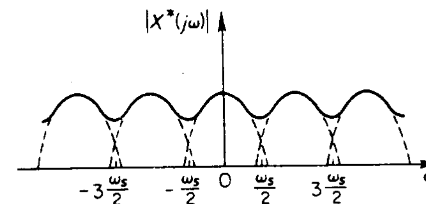


Fig. 13.7 Gráficos de $|X^*(j\omega)|$ em função de ω.

A Fig. 13.7 mostra gráficos de $|X^*(j\omega)|$ em função de ω para dois valores de T. Cada gráfico de $|X^*(j\omega)|$ *versus* ω consiste em $|X(j\omega)|$ repetido a cada $\omega_s = 2\pi/T$ rad/s. No espectro de freqüência, a componente $|X(j\omega)|/T$ é chamada de componente primária, e as outras componentes $|X[j(\omega \pm \omega_s k)]|/T$ são chamadas de componentes complementares.

Se $\omega_s > 2\omega_1$ ou $T < (\pi/\omega_1)$, não haverá sobreposição de componentes de $|X^*(j\omega)|$. Portanto a forma original de $|X(j\omega)|$ é preservada pelo processo de amostragem.

Se $\omega_s < 2\omega_1$ ou $(\pi/\omega_1) < T$, então a forma original de $|X(j\omega)|$ não aparece mais no gráfico de $|X^*(j\omega)|$ em função de ω.

Portanto, vemos que o sinal contínuo $x(t)$ pode ser reproduzido a partir do sinal amostrado $x^*(t)$ através de filtragem se e apenas se $\omega_s > 2\omega_1$ ou $T < (\pi/\omega_1)$.

Portanto, se o período de amostragem T é menor que π/ω_1, então o sinal contínuo pode ser reconstruído usando-se um filtro passa-baixas após o sinal ter sido amostrado. Se o filtro passa-baixas tem uma característica tal que deixa passar apenas sinais com freqüências menores que ω_1, então podemos obter o espectro de freqüência na saída do filtro exatamente como $1/T$ vezes $|X(j\omega)|$.

Note que na maioria dos sistemas a dados amostrados ou de tempo discreto, o circuito que segue o amostrador tem a característica de um filtro passa-baixas. Isto reduz as componentes de alta freqüência, alisa o sinal amostrado, e reproduz a grosso modo o sinal contínuo original.

Características de resposta em freqüência de dispositivos seguradores de ordem zero. Para atenuar as componentes complementares introduzidas pelo amostrador, o sinal amostrado é normalmente passado por um circuito segurador, um filtro passa-baixas.

Se o sinal entre dois instantes consecutivos de amostragem é aproximado por um polinômio de ordem n, então o dispositivo segurador é chamado de segurador de ordem n. Neste livro, consideramos apenas dispositivos seguradores de ordem zero. Tais dispositivos aproximam o sinal de saída $x(t)$ por um polinômio de grau

zero, ou uma constante. Portanto,

$$x(t) = x(kT) \qquad \text{para } kT \leq t < (k+1)T$$

onde $k = 0, 1, 2, \ldots$. A função de transferência de um dispositivo segurador de ordem zero é

$$G_h(s) = \frac{1 - e^{-Ts}}{s}$$

Substituindo $s = j\omega$ nesta função de transferência, obtemos

$$G_h(j\omega) = \frac{1 - e^{-Tj\omega}}{j\omega}$$

$$= \frac{2e^{-Tj\omega/2}(e^{Tj\omega/2} - e^{-Tj\omega/2})}{2j\omega}$$

$$= T\frac{\text{sen}\left(\frac{\omega T}{2}\right)}{\frac{\omega T}{2}}e^{-Tj\omega/2}$$

Em termos da freqüência de amostragem $\omega_s = 2\pi/T$,

$$G_h(j\omega) = \frac{2\pi}{\omega_s}\frac{\text{sen}\left(\frac{\pi\omega}{\omega_s}\right)}{\frac{\pi\omega}{\omega_s}}e^{-j\pi(\omega/\omega_s)}$$

A Fig. 13.8 mostra a curva de resposta em freqüência do dispositivo segurador de ordem zero. Claramente, ele tem característica de um filtro passa-baixas.

Soma de convolução. Considere o sistema de tempo discreto visto na Fig. 13.9. Aqui a seqüência de impulsos $x^*(t)$ é a entrada para o processo de tempo contínuo cuja função de transferência é $G(s)$. A saída do processo é um sinal contínuo $y(t)$. Se na saída há um outro amostrador, que está sincronizado em fase com o amostrador

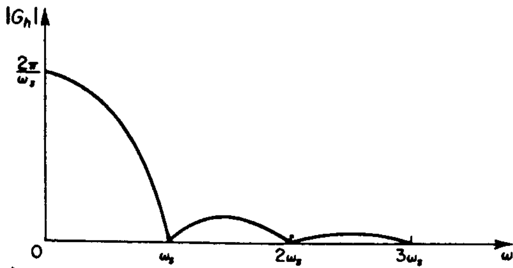


Fig. 13.8 Curva de resposta em freqüência do dispositivo segurador de ordem zero.

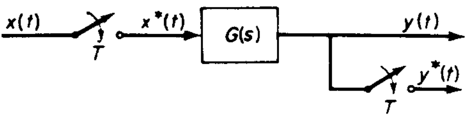


Fig. 13.9 Sistema de tempo discreto.

de entrada e opera no mesmo período de amostragem, então a saída é um trem de impulsos. Se $y(t) = 0$ para $t < 0$, então a transformada z de $y(t)$ é

$$\mathcal{Z}[y(t)] = Y(z) = \sum_{k=0}^{\infty} y(kT)z^{-k} \tag{13.13}$$

Na ausência de um amostrador na saída, se considerarmos um amostrador fictício (que é sincronizado em fase com o amostrador de entrada e opera à mesma taxa de amostragem) na saída e observarmos a seqüência de valores tomados por y apenas nos instantes $t = kT$, então a transformada z da saída pode também ser dada pela Eq. (13.13).

Deve-se notar que, ao realimentar o sinal de saída para a entrada, isto é, fechando a malha, a existência ou não existência do amostrador de saída dentro da malha faz uma diferença no comportamento do sistema. (Se o amostrador de saída está fora da malha, isto não fará nenhuma diferença na operação de malha-fechada.)

Para um processo linear estável invariante no tempo, sabe-se que a saída $y(t)$ do processo está relacionada com a entrada $x(t)$ através da integral de convolução, ou

$$y(t) = \int_0^t g(t-\tau)x(\tau)\,d\tau = \int_0^t x(t-\tau)g(\tau)\,d\tau$$

onde $g(t)$ é a função de resposta ao impulso do processo. Para sistemas de tempo discreto temos a soma de convolução, que é similar à integral de convolução. Como $x^*(t)$ é um trem de impulsos, a resposta do processo à entrada $x^*(t)$ é a soma das respostas a impulso individuais. Portanto, para $0 \leqslant t \leqslant kT$,

$$y(t) = g(t)x(0) + g(t-T)x(T) + g(t-2T)x(2T) + \cdots + g(t-kT)x(kT)$$

$$= \sum_{h=0}^{k} g(t-hT)x(hT)$$

Os valores da saída $y(t)$ nos instantes de amostragem $t = kT$ $(k = 0, 1, 2, \ldots)$ são dados por

$$y(kT) = \sum_{h=0}^{k} g(kT-hT)x(hT) \tag{13.14}$$

$$= \sum_{h=0}^{k} x(kT-hT)g(hT) \tag{13.15}$$

A soma na Eq. (13.14) ou (13.15) é chamada soma de convolução. Note que a notação simplificada

$$y(kT) = x(kT) * g(kT)$$

é freqüentemente chamada de soma de convolução.

Funções de transferência pulsada. Inicialmente notaremos que $y(kT)$ pode também ser escrito como

$$y(kT) = \sum_{h=0}^{\infty} g(kT - hT)x(hT) \qquad (k = 0, 1, 2, \ldots)$$

pois $g(kT - hT) = 0$ para $h > k$. Portanto,

$$\begin{aligned}
Y(z) &= \sum_{k=0}^{\infty} y(kT)z^{-k} \\
&= \sum_{k=0}^{\infty} \sum_{h=0}^{\infty} g(kT - hT)x(hT)z^{-k} \\
&= \sum_{m=0}^{\infty} \sum_{h=0}^{\infty} g(mT)x(hT)z^{-(m+h)} \\
&= \sum_{m=0}^{\infty} g(mT)z^{-m} \sum_{h=0}^{\infty} x(hT)z^{-h} \\
&= \sum_{m=0}^{\infty} g(mT)z^{-m}X(z) \\
&= G(z)X(z) \qquad (13.16)
\end{aligned}$$

onde

$$\begin{aligned}
G(z) &= \sum_{k=0}^{\infty} g(kT)z^{-k} \\
&= g(0) + g(T)z^{-1} + g(2T)z^{-2} + \cdots \\
&= z \text{ transformada } z \text{ de } g(t)
\end{aligned}$$

A Eq. (13.16) relaciona a saída pulsada do sistema com sua entrada em pulsos. A função $G(z)$, onde

$$G(z) = \frac{Y(z)}{X(z)}$$

é chamada de função de transferência z, ou função de transferência pulsada, do sistema de tempo discreto. A Fig.13.10 mostra um diagrama de bloco para uma função de transferência pulsada $G(z)$, juntamente com a entrada $X(z)$ e a saída $Y(z)$.

Procedimento geral para obtenção de funções de transferência pulsada. A função de transferência pulsada de um sistema pode ser obtida usando-se o seguinte procedimento:

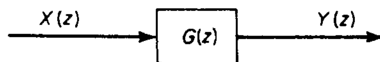


Fig. 13.10 Diagrama de bloco para um sistema de função de transferência pulsada.

1. Obtenha a função de transferência $G(s)$ do sistema.
2. Obtenha a função de resposta ao impulso $g(t)$, onde $g(t) = \mathcal{L}^{-1}[G(s)]$.
3. Calcule

$$G(z) = \sum_{k=0}^{\infty} g(kT)z^{-k}$$

onde $g(kT)$ é obtido de $g(t)$ substituindo t por kT. Para sistemas estáveis, a série infinita converge.

É importante lembrar que na abordagem por transformada z, ou a abordagem por função de transferência pulsada, para a análise de sistemas de tempo discreto, supõe-se que o sinal amostrado é um trem de impulsos cujas magnitudes, ou áreas, são iguais ao sinal de tempo contínuo nos instantes de amostragem. Tal suposição é válida apenas se a duração de amostragem do dispositivo amostrador é pequena quando comparada com a maior constante de tempo do sistema.

Exemplo 13.10 Obtenha a função de transferência pulsada do sistema visto na Fig. 13.11.

A função de transferência $G(s)$ pode ser expandida em frações parciais como segue:

$$\begin{aligned}
G(s) &= \frac{K}{(s+a)(s+b)} \\
&= \frac{K}{b-a}\left(\frac{1}{s+a} - \frac{1}{s+b}\right)
\end{aligned}$$

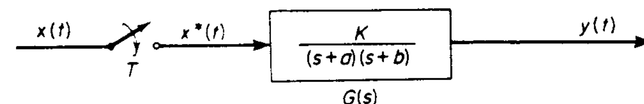


Fig. 13.11 Sistema de tempo discreto.

A função de resposta ao impulso é então obtida como

$$g(t) = \frac{K}{b-a}(e^{-at} - e^{-bt})$$

Portanto

$$g(kT) = \frac{K}{b-a}(e^{-akT} - e^{-bkT})$$

Então $G(z)$ é obtida como

$$G(z) = \sum_{k=0}^{\infty} \frac{K}{b-a}(e^{-akT} - e^{-bkT})z^{-k} \qquad (13.17)$$

Como

$$\sum_{k=0}^{\infty} e^{-akT}z^{-k} = 1 + \sum_{k=1}^{\infty} e^{-akT}z^{-k}$$
$$= 1 + \sum_{k=1}^{\infty} e^{-aT}z^{-1}(e^{-akT+aT}z^{-k+1})$$
$$= 1 + e^{-aT}z^{-1} \sum_{k=0}^{\infty} e^{-akT}z^{-k}$$

obtemos

$$\sum_{k=0}^{\infty} e^{-akT}z^{-k} = \frac{1}{1 - e^{-aT}z^{-1}} \tag{13.18}$$

Substituindo a Eq. (13.18) na Eq. (13.17), obtemos a seguinte função de transferência pulsada $G(z)$:

$$G(z) = \frac{K}{b-a}\left(\frac{1}{1 - e^{-aT}z^{-1}} - \frac{1}{1 - e^{-bT}z^{-1}}\right)$$
$$= \frac{K}{b-a}\frac{z(e^{-aT} - e^{-bT})}{(z - e^{-aT})(z - e^{-bT})}$$

Exemplo 13.11 Obtenha a função de transferência pulsada do sistema visto na Fig. 13.12. Do diagrama de bloco, obtemos

$$G(s) = \frac{C(s)}{M^*(s)} = \frac{1 - e^{-s}}{s^2(s+1)}$$
$$= (1 - e^{-s})\left(\frac{1}{s^2} - \frac{1}{s} + \frac{1}{s+1}\right)$$

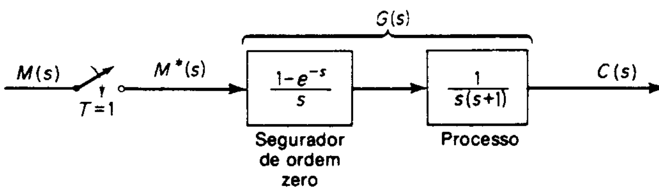


Fig. 13.12 Sistema de tempo discreto.

Portanto

$$g(t) = (t - 1 + e^{-t})1(t) - (t - 1 - 1 + e^{-(t-1)})1(t-1)$$

Como $T = 1$, temos $kT = k$ e

$$g(k) = (k - 1 + e^{-k}) - (k - 2 + e^{-(k-1)})$$
$$= e^{-k} + 1 - e^{-(k-1)} \qquad (k = 1, 2, 3, \ldots)$$
$$g(0) = 0$$

Portanto $G(z)$ é obtida como

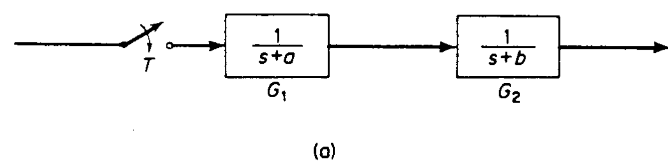
$$G(z) = \sum_{k=0}^{\infty} g(k)z^{-k}$$
$$= \sum_{k=0}^{\infty} (e^{-k} + 1 - e^{-(k-1)})z^{-k} + e - 2$$
$$= \frac{1-e}{1 - e^{-1}z^{-1}} + \frac{1}{1 - z^{-1}} + e - 2$$
$$= \frac{e^{-1}z^{-1} + (1 - 2e^{-1})z^{-2}}{1 - (1 + e^{-1})z^{-1} + e^{-1}z^{-2}}$$
$$= \frac{e^{-1}z + 1 - 2e^{-1}}{z^2 - (1 + e^{-1})z + e^{-1}}$$

Funções de transferência pulsada de elementos em cascata. É importante notar que as funções de transferência pulsada dos sistemas vistos nas Figs. 13.13(a) e (b) são diferentes. Para o sistema visto na Fig. 13.13(a), a função de transferência pulsada é, do Exemplo 13.10 (com $K = 1$),

$$G_1G_2(z) = \frac{z(e^{-aT} - e^{-bT})}{(b-a)(z - e^{-aT})(z - e^{-bT})}$$

Para o sistema visto na Fig. 13.13(b) (supondo que os dois amostradores estão sincronizados e têm o mesmo período de amostragem), obtemos

$$G_1(z)G_2(z) = \frac{z^2}{(z - e^{-aT})(z - e^{-bT})}$$



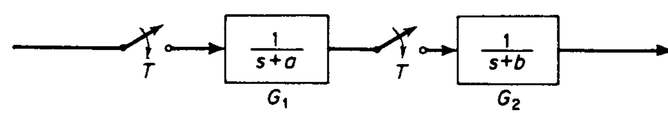


Fig. 13.13 (a) Sistema de tempo discreto com um amostrador; (b) sistema de tempo discreto com dois amostradores sincronizados.

Portanto

$$G_1G_2(z) \neq G_1(z)G_2(z)$$

Portanto, devemos ser cuidadosos e observar se há ou não um amostrador entre elementos em cascata.

Funções de transferência pulsada de sistemas a malha-fechada. Considere o sistema a malha-fechada da Fig. 13.14. Neste sistema, o erro atuante é amostrado. Do diagrama de bloco,

$$E(s) = R(s) - H(s)C(s)$$
$$C(s) = G(s)E^*(s)$$

Portanto

$$E(s) = R(s) - G(s)H(s)E^*(s)$$

Então, obtemos

$$E^*(s) = R^*(s) - GH^*(s)E^*(s)$$

ou

$$E^*(s) = \frac{R^*(s)}{1 + GH^*(s)}$$

Como

$$C^*(s) = G^*(s)E^*(s)$$

obtemos

$$C^*(s) = \frac{G^*(s)R^*(s)}{1 + GH^*(s)}$$

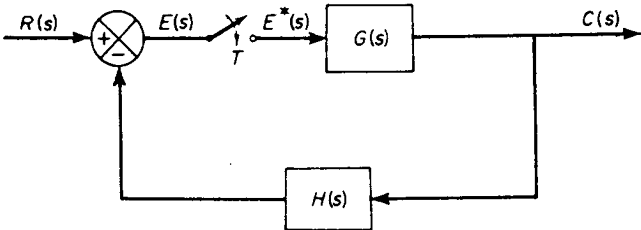


Fig. 13.14 Sistema de tempo discreto a malha-fechada.

Em termos da transformada z, $C(z)$ é dada por

$$C(z) = \frac{G(z)R(z)}{1 + GH(z)} \tag{13.19}$$

A transformada z inversa da Eq. (13.19) nos dá os valores da saída nos instantes de amostragem. A função de transferência pulsada do presente sistema a malha-

Tabela 13.3 Configurações típicas de sistemas de tempo discreto a malha-fechada e as correspondentes saídas $C(z)$

R(s), G(s), C(s), C(z), H(s)	$C(z) = \frac{G(z)R(z)}{1 + GH(z)}$
R(s), G(s), C(s), C(z), H(s)	$C(z) = \frac{G(z)R(z)}{1 + G(z)H(z)}$
R(s), G(s), C(s), C(z), H(s)	$C(z) = \frac{RG(z)}{1 + HG(z)}$
R(s), $G_1(s)$, $G_2(s)$, C(s), C(z), H(s)	$C(z) = \frac{G_2(z)RG_1(z)}{1 + G_1G_2H(z)}$
R(s), $G_1(s)$, $G_2(s)$, C(s), C(z), H(s)	$C(z) = \frac{G_1(z)G_2(z)R(z)}{1 + G_1(z)G_2H(z)}$

echada é

$$\frac{C(z)}{R(z)} = \frac{G(z)}{1 + GH(z)} \tag{13.20}$$

A Tabela 13.3 mostra cinco configurações típicas de sistemas a malha-fechada de tempo discreto. Para cada configuração, é vista a correspondente saída $C(z)$.

Exemplo 13.12 Obtenha a resposta a degrau unitário do sistema visto na Fig. 13.15.

A função de transferência pulsada deste sistema a malha-fechada é

$$\frac{C(z)}{R(z)} = \frac{G(z)}{1 + G(z)}$$

Como

$$G(s) = \frac{1 - e^{-s}}{s^2(s+1)}$$

referindo-se ao Exemplo 13.11, obtemos

$$G(z) = \frac{e^{-1}z + 1 - 2e^{-1}}{z^2 - (1 + e^{-1})z + e^{-1}}$$
$$= \frac{0{,}368z + 0{,}264}{z^2 - 1{,}368z + 0{,}368}$$

Portanto

$$\frac{C(z)}{R(z)} = \frac{0{,}368z + 0{,}264}{z^2 - z + 0{,}632}$$

Para uma entrada degrau unitário,

$$R(z) = \frac{z}{z - 1}$$

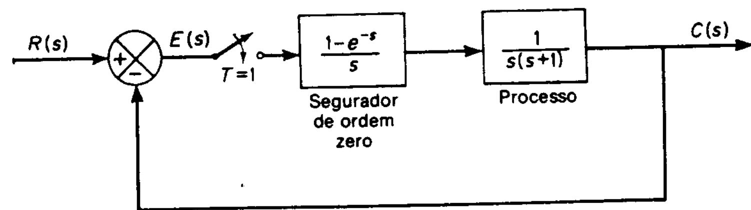


Fig. 13.15 Sistema de tempo discreto a malha-fechada.

A saída $C(z)$ é então obtida como segue:

$$C(z) = \frac{(0{,}368z + 0{,}264)z}{(z^2 - z + 0{,}632)(z - 1)}$$
$$= \frac{0{,}368z^2 + 0{,}264z}{z^3 - 2z^2 + 1{,}632z - 0{,}632}$$
$$= \frac{0{,}368z^{-1} + 0{,}264z^{-2}}{1 - 2z^{-1} + 1{,}632z^{-2} - 0{,}632z^{-3}}$$
$$= 0{,}368z^{-1} - z^{-2} + 1{,}4z^{-3} + 1{,}4z^{-4}$$
$$+ 1{,}147z^{-5} + 0{,}895z^{-6} - 0{,}802z^{-7} + \cdots$$

A transformada z inversa de $C(z)$ nos dá

$$c(0) = 0$$
$$c(1) = 0{,}368$$
$$c(2) = 1$$
$$c(3) = 1{,}4$$
$$c(4) = 1{,}4$$
$$c(5) = 1{,}147$$
$$c(6) = 0{,}895$$
$$c(7) = 0{,}802$$
$$\cdots$$

A saída $c(k)$ é mostrada no gráfico da Fig. 13.16.

A análise por transformada z não dará informação sobre a resposta entre instantes de amostragem. Portanto, a curva suave conectando os pontos é apenas aproximada.

Se desejamos obter informação sobre a resposta entre instantes de amostragem usando o método da transformada z, necessitamos modificar este método. O método da transformada z modificada, o método da amostragem submúltipla etc. nos permitem obter a resposta do sistema entre instantes de amostragem. (O leitor interessado em tais métodos modificados deve consultar livros especializados sobre sistemas a dados amostrados.)

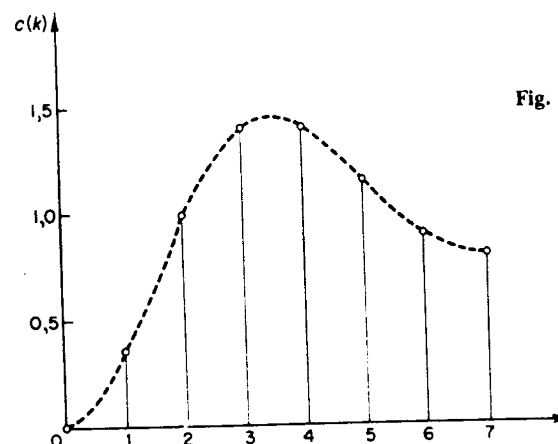


Fig. 13.16 Gráfico de $c(k)$ *versus* k.

13.6 ANÁLISE DE ESTABILIDADE NO PLANO z

Esta seção apresenta a análise de estabilidade de sistemas de tempo discreto no plano z. A análise de estabilidade de tais sistemas usando-se o segundo método de Liapunov será discutida na Seção 15.4.

Mapeamento do semiplano esquerdo no plano s sobre o plano z. Um sistema dinâmico linear é estável se todos os pólos da função de transferência estão no semiplano esquerdo do plano s. No plano z, o semiplano esquerdo em s corresponde ao círculo unitário centrado na origem, ou o semiplano esquerdo em s mapeia-se dentro do círculo unitário no plano z. Isto pode ser provado facilmente. Como

$$z = e^{Ts}, \qquad s = \sigma + j\omega$$

obtemos

$$|z| = e^{T\sigma}, \qquad \angle z = \omega T$$

No semiplano esquerdo em s, $\sigma < 0$. Portanto, o módulo de z varia entre 0 e 1. O eixo imaginário, ou $\sigma = 0$, corresponde ao círculo unitário no plano z. O interior do círculo corresponde ao semiplano esquerdo em s.

Note que como $\angle z = \omega T$, o ângulo de z varia de $-\infty$ a ∞ quando ω varia de $-\infty$ a ∞. Considere um ponto representativo no eixo $j\omega$ no plano s. À medida que este ponto se move de $-\pi/T$ a π/T no eixo $j\omega$, temos $z = 1$, e $\angle z$ varia de $-\pi$ a π na direção anti-horária no plano z. Quando o ponto representativo se move de π/T a $3\pi/T$ no eixo $j\omega$, o ponto correspondente no plano z descreve uma vez a circunferência unitária no sentido anti-horário. Portanto, quando o ponto no plano s se move de $-\infty$ a ∞ no eixo $j\omega$, percorremos a circunferência unitária no plano z um número infinito de vezes.

Desta análise, é claro que cada faixa de largura $2\pi/T$ no semiplano esquerdo do plano s é mapeada dentro do círculo unitário no plano z, como visto na Fig. 13.17(a). As Figs. 13.17(b) e (c) mostram regiões correspondentes nos planos s e z.

Análise de estabilidade. Discutiremos agora a estabilidade do sistema a malha-fechada visto na Fig. 13.14 no plano z. A relação saída-entrada do sistema a erro amostrado é dada pela Eq. (13.20), reescrita como

$$\frac{C(z)}{R(z)} = \frac{G(z)}{1 + GH(z)}$$

A estabilidade de tal sistema pode ser determinada pela localização das raízes da equação característica

$$1 + GH(z) = 0 \tag{13.21}$$

Como este polinômio em z pode ser convertido em uma razão de dois polinômios em z, para haver estabilidade todas as raízes z_i da equação característica, a Eq. (13.21), devem estar dentro do círculo unitário, ou

$$|z_i| < 1$$

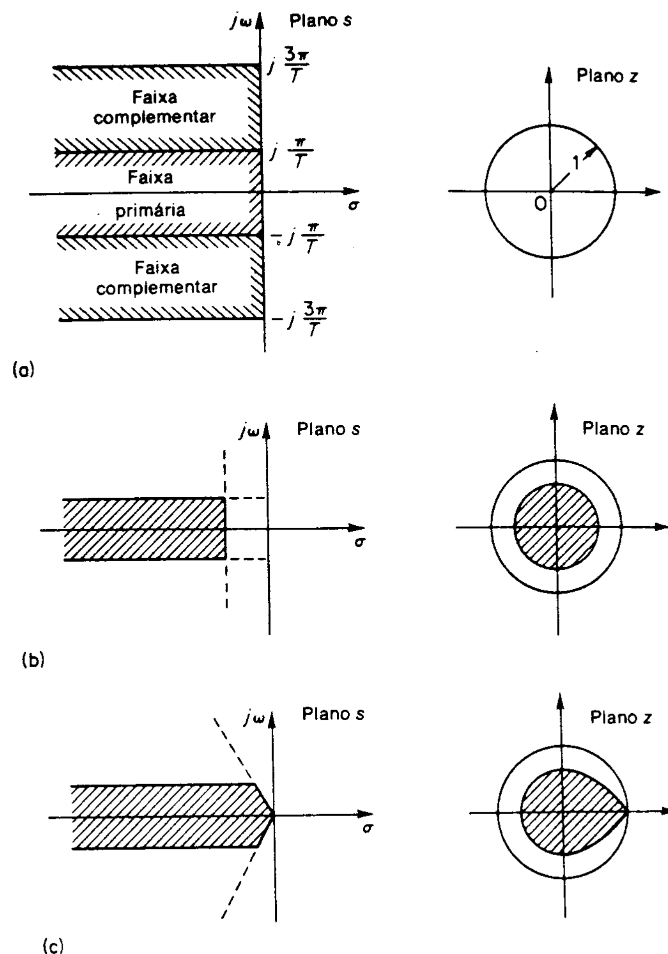


Fig. 13.17 (a) Mapeamento de faixas do plano s sobre o plano z; (b) e (c) regiões do plano s e as correspondentes regiões no plano z.

O sistema a malha-fechada $C(z)/R(z)$ se torna instável se qualquer pólo de malha-fechada estiver fora do círculo unitário e/ou qualquer pólo múltiplo estiver sobre a circunferência unitária.

Alguns métodos são disponíveis para a determinação da existência de raízes sobre ou fora do círculo unitário de um polinômio em z. Um método é modificar o critério de estabilidade de Routh. O critério de estabilidade de Routh nos diz se qualquer das raízes de um polinômio fica do lado direito do plano complexo. Como a seguinte transformação:

$$z = \frac{r + 1}{r - 1}$$

mapeia o interior do círculo unitário no plano z sobre o lado esquerdo do plano r, com esta transformação, o critério de estabilidade de Routh pode ser aplicado ao polinômio em r da mesma forma que em sistemas de tempo contínuo. O Exemplo 13.13 ilustra este método. Uma outra abordagem é aplicar o critério de estabilidade de Schur-Cohn. (Para uma discussão deste critério, consultar um livro especializado em sistemas a dados amostrados.)

Alternativamente, escrevendo uma equação de espaço de estados para o sistema de tempo discreto, podemos aplicar métodos de espaço de estados para determinar a estabilidade. Discutiremos este assunto nos Caps. 14 e 15.

Exemplo 13.13 Considere o sistema de tempo discreto visto na Fig. 13.18. A função de transferência de malha-aberta do sistema é

$$G(s) = \frac{10}{s(s+1)}$$

A transformada z de $G(s)$ é

$$G(z) = \frac{10(1-e^{-1})z}{(z-1)(z-e^{-1})}$$

A equação característica

$$1 + G(z) = 0$$

se torna

$$(z-1)(z-e^{-1}) + 10(1-e^{-1})z = 0$$

Observando-se que $e^{-1} = 0{,}368$, podemos simplificar a equação característica para

$$z^2 + 4{,}952z + 0{,}368 = 0 \qquad (13.22)$$

da qual podemos achar as raízes

$$z = -0{,}076, \qquad z = -4{,}876$$

Portanto, uma raiz da equação característica tem módulo maior do que a unidade, e o sistema é instável.

É importante notar que na ausência de um amostrador, um sistema de segunda-ordem sempre é estável. Na presença de um amostrador, entretanto, um sistema de segunda-ordem como este pode-se tornar instável para valores grandes do ganho. De fato, pode-se observar

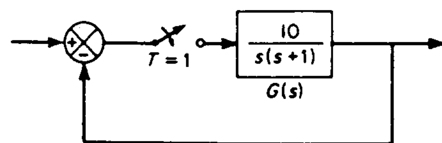


Fig. 13.18 Sistema de tempo discreto a malha-fechada.

que o sistema de segunda-ordem visto na Fig. 13.18 com

$$G(s) = \frac{K}{s(s+1)}$$

é estável apenas para $0 < K < 4{,}32$. (Veja Problema A.13.14.)

No presente sistema, a equação característica é de forma quadrática e pode ser resolvida facilmente. Para uma equação característica de grau superior, não é fácil determinar as raízes. Em tal caso, é conveniente transformar a equação característica em z em um polinômio em r usando a transformação

$$z = \frac{r+1}{r-1} \qquad (13.23)$$

e então aplicar o critério de estabilidade de Routh para o polinômio resultante.

Ilustraremos este método usando o sistema presente. Usando a transformação dada pela Eq. (13.23) podemos escrever a equação característica, Eq. (13.22), como

$$\left(\frac{r+1}{r-1}\right)^2 + 4{,}952\left(\frac{r+1}{r-1}\right) + 0{,}368 = 0$$

ou

$$6{,}32r^2 + 1{,}264r - 3{,}584 = 0$$

A tabela de Routh se torna

r^2	6,32	−3,584
r^1	1,264	0
r^0	−3,584	

Há uma mudança de sinal na primeira coluna da tabela de Routh. Portanto, há uma raiz no semiplano direito do plano r, o que implica que há uma raiz fora do círculo unitário no plano z. Isto corresponde ao resultado obtido previamente.

Note que, em geral, a estabilidade de um sistema de tempo discreto é melhorada quando o período de amostragem é diminuído, porque o aumento da taxa de amostragem tende a fazer o sistema se comportar de forma mais parecida com o sistema de tempo contínuo correspondente. (No presente exemplo, o sistema de tempo contínuo correspondente — um sistema de segunda-ordem — é estável para todos os valores positivos de K.)

Problemas Ilustrativos e Soluções

Problema A.13.1 Obtenha as transformadas z de a^k e $\mathbf{A}^k$, onde $\mathbf{A}$ é uma matriz $n \times n$.

Solução. Por definição, a transformada z de a^k é

$$\mathcal{Z}[a^k] = \sum_{k=0}^{\infty} a^k z^{-k}$$
$$= 1 + az^{-1} + a^2z^{-2} + \cdots$$
$$= \frac{1}{1 - az^{-1}}$$
$$= \frac{z}{z-a}$$

De forma semelhante,

$$\mathcal{Z}[\mathbf{A}^k] = \sum_{k=0}^{\infty} \mathbf{A}^k z^{-k}$$
$$= \mathbf{I} + \mathbf{A}z^{-1} + \mathbf{A}^2 z^{-2} + \cdots$$
$$= (\mathbf{I} - \mathbf{A}z^{-1})^{-1}$$
$$= (z\mathbf{I} - \mathbf{A})^{-1} z$$

Note que $\mathbf{A}^k$ pode ser obtida tomando-se a transformada z inversa de $(z\mathbf{I} - \mathbf{A})^{-1}z$, ou

$$\mathbf{A}^k = \mathcal{Z}^{-1}[(z\mathbf{I} - \mathbf{A})^{-1} z]$$

Problema A.13.2 Considere o sistema visto na Fig. 13.19. Derive a equação de diferenças que descreve a dinâmica do sistema quando a tensão de entrada aplicada é constante por trechos, ou

$$e(t) = e(kT) \quad \text{para } kT \le t < (k+1)T$$

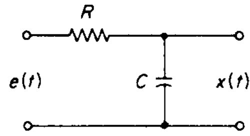


Fig. 13.19 Sistema a circuito elétrico.

Solução. Da Fig. 13.19, obtemos

$$RC\dot{x}(t) + x(t) = e(kT) \quad \text{para } kT \le t < (k+1)T$$

Tomando a transformada de Laplace desta última equação, considerando $t = kT$ como o instante inicial, obtemos

$$RC[sX(s) - x(kT)] + X(s) = \frac{e(kT)}{s}$$

ou

$$X(s) = \frac{e(kT)}{s} + \frac{x(kT) - e(kT)}{s + \frac{1}{RC}}$$

A transformada inversa de Laplace de $X(s)$ nos dá

$$x(t) = e(kT) + [x(kT) - e(kT)]e^{-(t-kT)/RC} \quad \text{para } kT \le t < (k+1)T$$

Substituindo $t = (k+1)T - 0$ nesta última equação e notando que $x(t)$ é uma função contínua do tempo e portanto $x((k+1)T - 0) = x((k+1)T + 0) = x((k+1)T)$, obtemos a equação de diferenças desejada, ou

$$x((k+1)T) = e^{-T/RC} x(kT) + (1 - e^{-T/RC})e(kT)$$

Problema A.13.3 Dado

$$X(z) = \frac{2z(z^2 - 1)}{(z^2 + 1)^2}$$

obtenha $x(kT)$.

Solução. Expandindo $X(z)/z$ em frações parciais, obtemos

$$\frac{X(z)}{z} = \frac{1}{(z+j)^2} + \frac{0}{z+j} + \frac{1}{(z-j)^2} + \frac{0}{z-j}$$

Portanto

$$X(z) = \frac{z}{(z+j)^2} + \frac{z}{(z-j)^2}$$
$$= (e^{j(\pi/2)})\frac{ze^{-j(\pi/2)}}{(z - e^{-j(\pi/2)})^2} + (e^{-j(\pi/2)})\frac{ze^{j(\pi/2)}}{(z - e^{j(\pi/2)})^2}$$

Usando a Tabela 13.1, obtemos a transformada z inversa de $X(z)$ como segue

$$x(kT) = e^{j(\pi/2)}ke^{-j(\pi/2)k} + e^{-j(\pi/2)}ke^{j(\pi/2)k}$$
$$= jk\left[\cos\frac{\pi k}{2} - j\,\text{sen}\frac{\pi k}{2}\right] - jk\left[\cos\frac{\pi k}{2} + j\,\text{sen}\frac{\pi k}{2}\right]$$
$$= 2k\,\text{sen}\frac{\pi k}{2} \qquad (k = 0, 1, 2, \ldots)$$

Problema A.13.4 Mostre que a transformada de Laplace de

$$x^*(t) = \sum_{k=0}^{\infty} x(t)\,\delta(t - kT)$$

é

$$X^*(s) = \sum_{k=0}^{\infty} x(kT)e^{-kTs}$$
$$= X(s) * \frac{1}{1 - e^{-Ts}}$$
$$= \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} X(p)\frac{1}{1 - e^{-T(s-p)}}\,dp$$

onde c é a abscissa de convergência para $X(s) = \mathcal{L}[x(t)]$.

Solução. Inicialmente, notemos que

$$\mathcal{L}[\delta(t - kT)] = e^{-kTs}$$

Portanto

$$X^*(s) = \mathcal{L}[x^*(t)] = \mathcal{L}[\sum_{k=0}^{\infty} x(t)\,\delta(t - kT)]$$
$$= \mathcal{L}[\sum_{k=0}^{\infty} x(kT)\delta(t - kT)]$$
$$= \sum_{k=0}^{\infty} x(kT)e^{-kTs}$$

Como

$$\mathscr{L}[\sum_{k=0}^{\infty}\delta(t-kT)] = 1 + e^{-Ts} + e^{-2Ts} + \cdots$$
$$= \frac{1}{1-e^{-Ts}}$$

obtemos

$$X^*(s) = \mathscr{L}[x^*(t)] = \mathscr{L}[\sum_{k=0}^{\infty} x(t)\,\delta(t-kT)]$$
$$= \mathscr{L}[x(t)\sum_{k=0}^{\infty}\delta(t-kT)]$$
$$= X(s) * \frac{1}{1-e^{-Ts}}$$

A transformada de Laplace do produto de duas funções $f(t)$ e $g(t)$ é

$$\mathscr{L}[f(t)g(t)] = \int_0^{\infty} f(t)g(t)\,e^{-st}\,dt$$

A integral de inversão definida pela Eq. (2.8) é

$$f(t) = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} F(s)\,e^{st}\,ds \qquad (t > 0)$$

onde c é a abscissa de convergência para $F(s)$. Portanto,

$$\mathscr{L}[f(t)g(t)] = \frac{1}{2\pi j}\int_0^{\infty}\int_{c-j\infty}^{c+j\infty} F(p)\,e^{pt}\,dp\,g(t)\,e^{-st}\,dt$$

Em virtude da convergência uniforme das integrais consideradas, podemos inverter a ordem da integração, ou

$$\mathscr{L}[f(t)g(t)] = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} F(p)\,dp\int_0^{\infty} g(t)\,e^{-(s-p)t}\,dt$$

Notando que

$$\int_0^{\infty} g(t)\,e^{-(s-p)t}\,dt = G(s-p)$$

obtemos

$$\mathscr{L}[f(t)g(t)] = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} F(p)G(s-p)\,dp$$

Portanto

$$\mathscr{L}[x^*(t)] = \mathscr{L}[x(t)\sum_{k=0}^{\infty}\delta(t-kT)]$$
$$= \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} X(p)\frac{1}{1-e^{-T(s-p)}}\,dp$$

Problema A.13.5 Supondo que os pólos de $X(s)$ estão no semiplano esquerdo do plano s e que $X(s)$ pode ser expresso como uma razão de polinômios com um denominador de grau pelo menos 2 a mais que o do numerador [o que significa que $\lim_{s\to\infty} sX(s) = x(0+) = 0$], calcule

$$X^*(s) = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty}\frac{X(p)}{1-e^{-T(s-p)}}\,dp$$

onde o contorno é uma linha paralela ao eixo imaginário no plano p que separa os pólos de $X(p)$ daqueles de $1/[1-e^{-T(s-p)}]$, e mostre que

$$X(z) = \frac{1}{T}\sum_{k=-\infty}^{\infty} X\left(s + j\frac{2\pi}{T}k\right)\Big|_{s=(1/T)\ln z}$$

Solução. Os pólos de $1/[1-e^{-T(s-p)}]$ podem ser obtidos resolvendo-se

$$1 - e^{-T(s-p)} = 0$$

e então

$$p = s \pm j\frac{2\pi}{T}k \qquad (k = 0, 1, 2, \ldots)$$

Portanto há um número infinito de pólos.

Para calcular a integral dada, vamos escolher o contorno que consiste na reta de $c - j\infty$ a $c + j\infty$ e o semicírculo Γ de um raio infinito no semiplano direito do plano p, como visto na Fig. 13.20. O contorno fechado engloba todos os pólos de $1/[1-e^{-T(s-p)}]$, mas não engloba quaisquer pólos de $X(p)$.

Agora $X^*(s)$ pode ser escrita como

$$X^*(s) = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty}\frac{X(p)}{1-e^{-T(s-p)}}\,dp$$
$$= \frac{1}{2\pi j}\oint\frac{X(p)}{1-e^{-T(s-p)}}\,dp - \frac{1}{2\pi j}\int_{\Gamma}\frac{X(p)}{1-e^{-T(s-p)}}\,dp$$

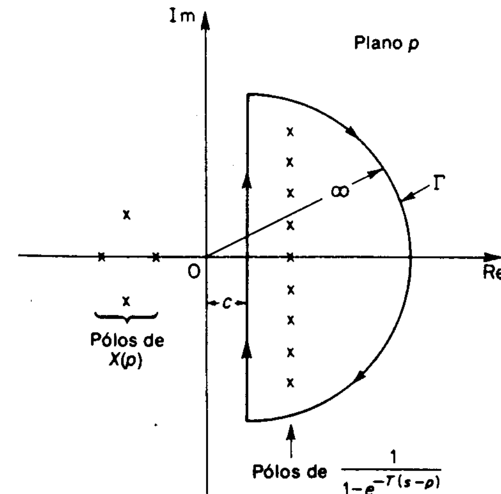


Fig. 13.20 Contorno no plano p.

Vamos considerar a integral ao longo do semicírculo Γ. Como o grau do denominador de $X(s)$ é pelo menos 2 a mais do que o grau do numerador, pode-se mostrar que o valor desta integral é zero, ou

$$\frac{1}{2\pi j}\int_{\Gamma}\frac{X(p)}{1-e^{-T(s-p)}}\,dp = 0$$

Portanto,

$$X^*(s) = \frac{1}{2\pi j}\oint\frac{X(p)}{1-e^{-T(s-p)}}\,dp$$

A integral ao longo do contorno fechado pode ser obtida calculando-se os resíduos. Portanto,

$$X^*(s) = -\sum_{k=-\infty}^{\infty}\left.\frac{X(p)}{\dfrac{d}{dp}[1-e^{-T(s-p)}]}\right|_{p=s+j(2\pi/T)k}$$
$$= \frac{1}{T}\sum_{k=-\infty}^{\infty}X\left(s+j\frac{2\pi}{T}k\right)$$

Portanto,

$$X(z) = \frac{1}{T}\sum_{k=-\infty}^{\infty}X\left(s+j\frac{2\pi}{T}k\right)\bigg|_{s=(1/T)\ln z}$$

Problema A.13.6 Supondo que os pólos de $X(s)$, uma razão de dois polinômios em s, estão no semiplano esquerdo de s e que $\lim_{s\to\infty} X(s) = 0$, calcule

$$X^*(s) = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty}\frac{X(p)}{1-e^{-T(s-p)}}\,dp$$

e mostre que

$$X(z) = \sum_{i=1}^{m}\left\{\frac{1}{(n_i-1)!}\frac{d^{n_i-1}}{ds^{n_i-1}}\left[\frac{(s-s_i)^{n_i}X(s)z}{z-e^{Ts}}\right]\right\}_{s=s_i}$$

onde os s_i são os pólos de $X(s)$ e n_i é a ordem do pólo em $s = s_i$, ou

$$X(z) = \sum\left[\text{resíduos de } \frac{X(s)z}{z-e^{Ts}} \text{ nos pólos de } X(s)\right]$$

Solução. Como obtido no Problema A.13.5, os pólos de $1/[1-e^{-T(s-p)}]$ no plano p são

$$p = s \pm j\frac{2\pi}{T}k \qquad (k = 0, 1, 2, \ldots)$$

Para calcular a dada integral, vamos escrever

$$X^*(s) = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty}\frac{X(p)}{1-e^{-T(s-p)}}\,dp$$
$$= \frac{1}{2\pi j}\oint\frac{X(p)}{1-e^{-T(s-p)}}\,dp - \frac{1}{2\pi j}\int_{\Gamma}\frac{X(p)}{1-e^{-T(s-p)}}\,dp$$

O contorno fechado consiste na reta de $c-j\infty$ a $c+j\infty$ e o semicírculo Γ de raio infinito no lado esquerdo do plano p, como visto na Fig. 13.21. No caso presente determinamos que a integral ao longo do semicírculo infinito Γ não é sempre zero. Se não for zero, considere $X(p)$ como $\lim_{\varepsilon\to 0} e^{\varepsilon p}X(p)$, $(\varepsilon > 0)$. Então podemos mostrar que a integral ao longo de Γ é zero. Portanto,

$$\frac{1}{2\pi j}\int_{\Gamma}\frac{X(p)}{1-e^{-T(s-p)}}\,dp = 0$$

Portanto

$$X^*(s) = \frac{1}{2\pi j}\oint\frac{X(p)}{1-e^{-T(s-p)}}\,dp$$

Como os pólos de $X(p)$ são $p_1, p_2, \ldots p_m$,

$$X^*(s) = \sum\left[\text{resíduos de } \frac{X(p)}{1-e^{-T(s-p)}} \text{ nos pólos de } X(p)\right]$$
$$= \sum_{i=1}^{m}\frac{1}{(n_i-1)!}\frac{d^{n_i-1}}{dp^{n_i-1}}\left[(p-p_i)^{n_i}\frac{X(p)}{1-e^{-T(s-p)}}\right]_{p=p_i}$$

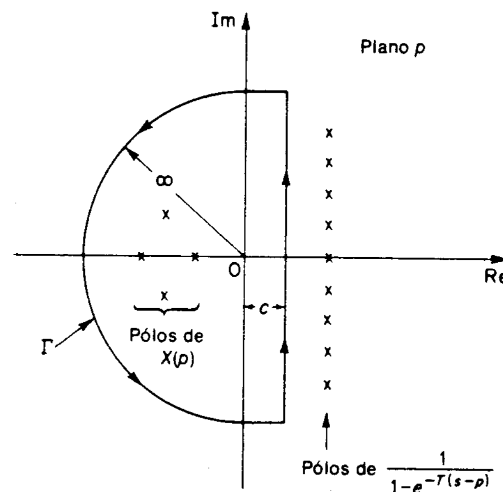


Fig. 13.21 Contorno no plano p.

Substituindo $z = e^{Ts}$ nesta última equação, obtemos

$$X(z) = \sum_{i=1}^{m}\frac{1}{(n_i-1)!}\frac{d^{n_i-1}}{dp^{n_i-1}}\left[(p-p_i)^{n_i}X(p)\frac{z}{z-e^{Tp}}\right]_{p=p_i}$$

Mudando a variável complexa de p para s, obtemos

$$X(z) = \sum_{i=1}^{m}\frac{1}{(n_i-1)!}\frac{d^{n_i-1}}{ds^{n_i-1}}\left[(s-s_i)^{n_i}X(s)\frac{z}{z-e^{Ts}}\right]_{s=s_i}$$
$$= \sum\left[\text{resíduos de } \frac{X(s)z}{z-e^{Ts}} \text{ nos pólos de } X(s)\right]$$

Problema A.13.7 Obtenha a transformada z de te^{-at}, onde $t \geqslant 0$. Note que $\mathscr{L}[te^{-at}] = 1/(s+a)^2$.

Solução

$$F(z) = \left[\text{resíduo de } \frac{X(s)z}{z - e^{Ts}} \text{ em } s = -a\right]$$
$$= \left[\text{resíduo de } \frac{z}{(s+a)^2(z - e^{Ts})} \text{ em } s = -a\right]$$
$$= \frac{d}{ds}\left[\frac{z}{z - e^{Ts}}\right]_{s=-a}$$
$$= \left.\frac{Tze^{Ts}}{(z - e^{Ts})^2}\right|_{s=-a}$$
$$= \frac{Tze^{-aT}}{(z - e^{-aT})^2}$$

Problema A.13.8 Dado

$$X(s) = \frac{s+3}{(s+1)(s+2)}$$

obtenha $X(z)$.

Solução

$$X(z) = \left[\text{resíduo de } \frac{(s+3)z}{(s+1)(s+2)(z - e^{Ts})} \text{ em } s = -1\right]$$
$$+ \left[\text{resíduo de } \frac{(s+3)z}{(s+1)(s+2)(z - e^{Ts})} \text{ em } s = -2\right]$$
$$= \left.\frac{(s+3)z}{(s+2)(z - e^{Ts})}\right|_{s=-1} + \left.\frac{(s+3)z}{(s+1)(z - e^{Ts})}\right|_{s=-2}$$
$$= \frac{2z}{z - e^{-T}} - \frac{z}{z - e^{-2T}}$$

Problema A.13.9 Prove que

$$\mathcal{Z}[\sum_{k=0}^{n} x(k)] = \frac{z}{z-1} X(z)$$
$$\mathcal{Z}[\sum_{k=0}^{n-1} x(k)] = \frac{1}{z-1} X(z)$$

Em seguida prove que

$$\sum_{k=0}^{\infty} x(k) = \lim_{z \to 1} X(z)$$

Solução. A transformada z da primeira diferença "para a frente" (*forward*) é

$$\mathcal{Z}[\Delta f(k)] = \mathcal{Z}[f(k+1)] - \mathcal{Z}[f(k)]$$
$$= [z F(z) - z f(0)] - F(z)$$
$$= (z-1) F(z) - z f(0)$$

Vamos escrever

$$x(n) = \sum_{k=0}^{n} x(k) - \sum_{k=0}^{n-1} x(k) = \Delta \sum_{k=0}^{n-1} x(k)$$

Então

$$\mathcal{Z}[\Delta \sum_{k=0}^{n-1} x(k)] = \mathcal{Z}[x(n)] = X(z) = (z-1)\, \mathcal{Z}[\sum_{k=0}^{n-1} x(k)]$$

como o valor da soma $\sum_{k=0}^{n-1} x(k)$ para $n = 0$ é igual a zero. Portanto,

$$\mathcal{Z}[\sum_{k=0}^{n-1} x(k)] = \frac{X(z)}{z-1}$$

e

$$\mathcal{Z}[\sum_{k=0}^{n} x(k)] = \mathcal{Z}[\sum_{k=0}^{n-1} x(k)] + \mathcal{Z}[x(n)] = \frac{X(z)}{z-1} + X(z)$$

ou

$$\mathcal{Z}[\sum_{k=0}^{n} x(k)] = \frac{z}{z-1} X(z)$$

Usando o teorema do valor final, achamos

$$\lim_{n \to \infty}\left[\sum_{k=0}^{n} x(k)\right] = \lim_{z \to 1}\left[(z-1)\frac{z}{z-1} X(z)\right]$$

ou

$$\sum_{k=0}^{\infty} x(k) = \lim_{z \to 1} X(z)$$

Problema A.13.10 Prove as seguintes relações. Note que $X(z) = \mathcal{Z}[x(t)]$.

1. $$\mathcal{Z}[a^k x(t)] = X\left(\frac{z}{a}\right)$$
2. $$\mathcal{Z}[e^{-at} x(t)] = X(ze^{aT})$$
3. $$\mathcal{Z}[t x(t)] = -Tz \frac{d}{dz} X(z)$$
4. $$\mathcal{Z}[t] = \frac{Tz}{(z-1)^2}$$

Solução

1. $$\mathcal{Z}[a^k x(t)] = \sum_{k=0}^{\infty} a^k x(kT) z^{-k}$$
$$= \sum_{k=0}^{\infty} x(kT)\left(\frac{z}{a}\right)^{-k}$$
$$= X\left(\frac{z}{a}\right)$$

2.
$$\mathcal{Z}[e^{-at}x(t)] = \sum_{k=0}^{\infty} e^{-akT}x(kT)z^{-k}$$
$$= \sum_{k=0}^{\infty} x(kT)(ze^{aT})^{-k}$$
$$= X(ze^{aT})$$

3.
$$\mathcal{Z}[tx(t)] = \sum_{k=0}^{\infty} kTx(kT)z^{-k}$$
$$= -T\sum_{k=0}^{\infty} x(kT)z\frac{d}{dz}(z^{-k})$$
$$= -Tz\frac{d}{dz}[\sum_{k=0}^{\infty} x(kT)z^{-k}]$$
$$= -Tz\frac{d}{dz}X(z)$$

4. Como $\mathcal{Z}[1(t)] = z/(z-1)$, referindo-se à relação que acabamos de obter, vemos que

$$\mathcal{Z}[t] = -Tz\frac{d}{dz}\{\mathcal{Z}[1(t)]\}$$
$$= -Tz\frac{d}{dz}\left(\frac{z}{z-1}\right)$$
$$= \frac{Tz}{(z-1)^2}$$

Problema A.13.11 Obtenha a solução da seguinte equação de diferenças em termos de $x(0)$ e $x(1)$:

$$x(k+2) + ax(k+1) + bx(k) = 0$$

onde a e b são constantes e $k \geq 0$.

Solução. As transformadas z de $x(k+2)$, $x(k+1)$, e $x(k)$ são dadas, respectivamente, por

$$\mathcal{Z}[x(k+2)] = z^2X(z) - z^2x(0) - zx(1)$$
$$\mathcal{Z}[x(k+1)] = zX(z) - zx(0)$$
$$\mathcal{Z}[x(k)] = X(z)$$

Portanto, a transformada z da equação de diferenças dada se torna

$$[z^2X(z) - z^2x(0) - zx(1)] + a[zX(z) - zx(0)] + bX(z) = 0$$

ou

$$(z^2 + az + b)X(z) = (z^2 + az)x(0) + zx(1)$$

Resolvendo para $X(z)$, temos

$$X(z) = \frac{(z^2 + az)x(0) + zx(1)}{z^2 + az + b}$$

Chamando as raízes do polinómio do denominador de m_1 e m_2, ou

$$z^2 + az + b = (z - m_1)(z - m_2)$$

e expandindo $X(z)/z$ em frações parciais, obtemos

$$\frac{X(z)}{z} = \frac{(m_1 - a)x(0) + x(1)}{m_1 - m_2}\frac{1}{z - m_1} + \frac{(m_2 + a)x(0) + x(1)}{m_2 - m_1}\frac{1}{z - m_2}$$

Notando que $a = -m_1 - m_2$, obtemos

$$X(z) = \frac{m_2x(0) - x(1)}{m_2 - m_1}\frac{z}{z - m_1} + \frac{m_1x(0) - x(1)}{m_1 - m_2}\frac{z}{z - m_2}$$

A transformada z inversa de $X(z)$ dá a solução da equação de diferenças fornecida. Ou seja,

$$x(k) = \frac{m_2x(0) - x(1)}{m_2 - m_1}m_1{}^k + \frac{m_1x(0) - x(1)}{m_1 - m_2}m_2{}^k \qquad (k = 0, 1, 2, \ldots)$$

Problema A.13.12 Obtenha a função de transferência pulsada do sistema a malha-fechada visto na Fig. 13.22.

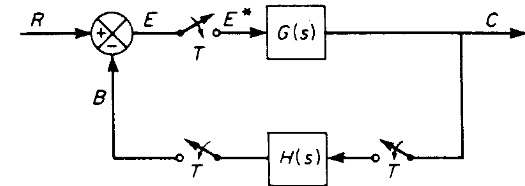


Fig. 13.22 Sistema de tempo discreto a malha-fechada.

Solução. Do diagrama de blocos da Fig. 13.22, obtemos

$$E(s) = R(s) - B(s)$$

ou

$$E(z) = R(z) - B(z)$$

Também

$$C(z) = G(z)E(z)$$
$$B(z) = H(z)C(z)$$

Portanto

$$C(z) = [R(z) - H(z)C(z)]G(z)$$

ou

$$C(z)[1 + H(z)G(z)] = R(z)G(z)$$

A função de transferência pulsada se torna

$$\frac{C(z)}{R(z)} = \frac{G(z)}{1 + H(z)G(z)}$$

Problema A.13.13 Obtenha a função de transferência pulsada do sistema a malha-fechada visto na Fig. 13.23.

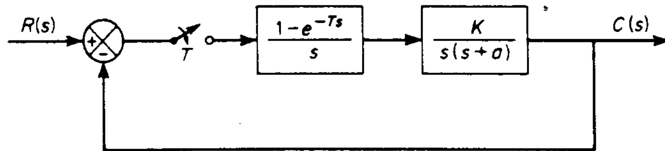


Fig. 13.23 Sistema de tempo discreto a malha-fechada.

Solução. A transformada z da função de transferência de malha-aberta é

$$\mathcal{Z}\left[\frac{K(1 - e^{-Ts})}{s^2(s + a)}\right] = K(1 - z^{-1})\mathcal{Z}\left[\frac{1}{s^2(s + a)}\right]$$

Note que

$$\frac{1}{s^2(s + a)} = \frac{1}{as^2} - \frac{1}{a^2 s} + \frac{1}{a^2(s + a)}$$

Da Tabela 13.1. obtemos a transformada z desta última equação como segue:

$$\mathcal{Z}\left[\frac{1}{s^2(s + a)}\right] = \frac{Tz}{a(z - 1)^2} - \frac{z}{a^2(z - 1)} + \frac{z}{a^2(z - e^{-aT})}$$

Portanto

$$\mathcal{Z}\left[\frac{K(1 - e^{-Ts})}{s^2(s + a)}\right] = \frac{KT}{a(z - 1)} - \frac{K}{a^2} + \frac{K(z - 1)}{a^2(z - e^{-aT})}$$
$$= \frac{K[(aT - 1 + e^{-aT})z + (1 - e^{-aT} - aTe^{-aT})]}{a^2(z - 1)(z - e^{-aT})}$$

A função de transferência de malha-fechada é então

$$\frac{C(z)}{R(z)} = \frac{K[(aT - 1 + e^{-aT})z + (1 - e^{-aT} - aTe^{-aT})]}{a^2z^2 + [K(aT - 1 + e^{-aT}) - a^2(1 + e^{-aT})]z + (1 - e^{-aT} - aTe^{-aT} + a^2e^{-aT})}$$

Problema A.13.14 Considere o sistema visto na Fig. 13.24. Mostre que o sistema de tempo discreto é estável se e apenas se

$$0 < K < 2\coth\left(\frac{T}{2T_1}\right)$$

Solução. Como

$$G(s) = \frac{K}{s(T_1 s + 1)}$$

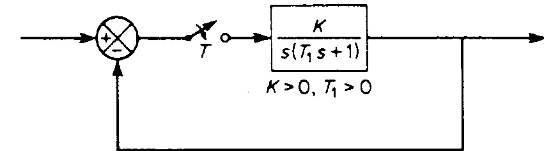


Fig. 13.24 Sistema de tempo discreto a malha-fechada.

obtemos

$$G(z) = \frac{K(1 - e^{-T/T_1})z}{(z - 1)(z - e^{-T/T_1})}$$

A equação característica é

$$z^2 + [K(1 - e^{-T/T_1}) - (1 + e^{-T/T_1})]z + e^{-T/T_1} = 0 \qquad (13.24)$$

Usando a seguinte transformação:

$$z = \frac{r + 1}{r - 1}$$

a Eq. (13.24) se torna

$$[K(1 - e^{-T/T_1})]r^2 + 2(1 - e^{-T/T_1})r + 2(1 + e^{-T/T_1}) - K(1 - e^{-T/T_1}) = 0$$

A tabela de Routh é

r^2	$K(1 - e^{-T/T_1})$	$2(1 + e^{-T/T_1}) - K(1 - e^{-T/T_1})$
r^1	$2(1 - e^{-T/T_1})$	0
r^0	$2(1 + e^{-T/T_1}) - K(1 - e^{-T/T_1})$	

Como $1 - e^{-T/T_1} > 0$, o sistema é estável se e apenas se

$$K > 0$$
$$\frac{2(1 + e^{-T/T_1})}{1 - e^{-T/T_1}} > K$$

Portanto o sistema é estável se e apenas se

$$0 < K < \frac{2(1 + e^{-T/T_1})}{1 - e^{-T/T_1}} = 2\coth\left(\frac{T}{2T_1}\right)$$

ROBLEMAS

roblema B.13.1 Obtenha a transformada z de $g(t - kT)$, onde $g(t)$ é a função de resposta ao npulso de um sistema linear.

roblema B.13.2 Se $X(z) = X_1(z)X_2(z)$, mostre que

$$x(kT) = \sum_{h=0}^{k} x_1(kT - hT)x_2(hT)$$

roblema B.13.3 Dado

$$X(z) = \frac{z}{(z-1)^2(z-2)}$$

etermine $x(kT)$.

roblema B.13.4 Dado

$$X(z) = \frac{z^2}{(ze-1)^3}$$

determine $x(kT)$ computando o resíduo de $X(z)z^{k-1}$ no pólo $z = 1/e$.

roblema B.13.5 Dado

$$X(z) = \frac{z(1-e^{-T})}{(z-1)(z-e^{-T})}$$

determine $x(kT)$ expandindo $X(z)$ em uma série convergente de potências em z^{-1}.

Problema B.13.6 Dado

$$G(s) = \frac{K}{s(s+a)}$$

determine a função de transferência pulsada $G(z)$.

Problema B.13.7 Dado

$$G(s) = \frac{\omega_0}{s^2 + \omega_0^2}$$

determine a função de transferência pulsada $G(z)$.

Problema B.13.8 Considere o circuito RC visto na Fig. 13.25. Determine $y(kT)$ quando a tensão de entrada é dada por $x(t) = 100e^{-t}$ volts.

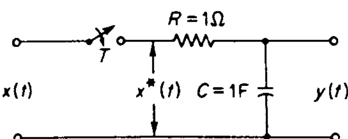


Fig. 13.25 Circuito RC com amostrador.

Problema B.13.9 Determine a solução da seguinte equação de diferenças:

$$x(k+2) + 2x(k+1) + x(k) = u(k), \qquad x(0) = 0, \qquad x(1) = 0$$

onde

$$u(k) = k \qquad (k = 0, 1, 2, \ldots)$$

14

Análise de Sistemas de Controle por Espaço de Estados

14.1 INTRODUÇÃO

Limitações da teoria convencional de controle. Nos Caps. 8 a 10, mostramos que o método do lugar das raízes e os métodos de resposta em freqüência são bastante úteis para tratar de sistemas de uma entrada e uma saída. Por exemplo, usando testes de resposta em freqüência de malha-aberta, podemos predizer o comportamento dinâmico do sistema a malha-fechada. Se necessário, o comportamento dinâmico de um sistema complexo pode ser melhorado inserindo-se um simples compensador avançador ou atrasador de fase. As técnicas da teoria convencional de controle são conceitualmente simples e requerem apenas uma pequena quantidade de computações.

Na teoria convencional de controle, apenas os sinais de entrada, saída e de erro são considerados importantes; a análise e projeto de sistemas de controle são feitos usando-se funções de transferência, juntamente com uma variedade de técnicas gráficas tais como o método do lugar das raízes e gráficos de Nyquist. A característica essencial da teoria de controle convencional é que ela é baseada na relação de entrada-saída do sistema, ou a função de transferência.

A principal desvantagem da teoria de controle convencional é que, de modo geral, ela é aplicável apenas para sistemas lineares invariantes no tempo tendo uma única entrada e uma única saída. Ela é impotente para sistemas variantes no tempo, sistemas não lineares (exceto os mais simples), e sistemas de múltiplas entradas e múltiplas saídas. Portanto as técnicas convencionais (os métodos do lugar das raízes e de resposta em freqüência) não se aplicam para o projeto de

sistemas ótimos ou de sistemas adaptativos, que são em geral variantes no tempo e/ou não lineares.

Uma nova abordagem para a análise e projeto de sistemas de controle — a teoria de controle moderno. A tendência atual em sistemas de engenharia é para maior complexidade, devido principalmente às necessidades de tarefas complexas e boa precisão. Sistemas complexos podem ter múltiplas entradas e múltiplas saídas e podem ser variantes no tempo. Em virtude da necessidade de satisfazer especificações cada vez mais rigorosas no desempenho de sistemas de controle, o aumento na complexidade do sistema e o fácil acesso a computadores, a teoria de controle moderno, que é uma nova abordagem na análise e projeto de sistemas de controle complexos, foi desenvolvida em torno de 1960. Esta nova abordagem é baseada no conceito de estado. O conceito de estado por si só não é novo, pois já existia faz tempo no campo da dinâmica clássica e outras áreas. (De fato, o método do plano de fase, discutido no Cap. 13, é um método de espaço de estados bidimensional.)

Teoria de controle moderno *versus* teoria de controle convencional. A teoria de controle moderno contrasta com a teoria de controle convencional, em que a primeira é aplicável para sistemas de múltiplas entradas e múltiplas saídas, que podem ser lineares ou não lineares, invariantes ou variantes no tempo, enquanto que a segunda só é aplicável para sistemas lineares invariantes no tempo de uma entrada e uma saída. Também, a teoria de controle moderno é essencialmente uma abordagem de domínio de tempo, enquanto que a teoria de controle convencional é uma abordagem de domínio da freqüência complexa.

O projeto de sistemas na teoria clássica de controle é baseado em procedimentos de tentativa e erro, que, em geral, não fornecem sistemas de controle ótimos. O projeto de sistemas na teoria moderna de controle, por outro lado, permite ao engenheiro projetar sistemas de controle ótimos com respeito a dados critérios de desempenho. Adicionalmente, o projeto na teoria de controle moderno pode ser feito para uma classe de entradas, ao invés de uma função de entrada específica, tal como a função impulsiva, função degrau, ou função senoidal. Também, a teoria de controle moderno permite ao engenheiro incluir condições iniciais no projeto.

Antes de prosseguir, devemos definir estado, variáveis de estado, vetor de estados, e espaço de estados.

Estado. O estado de um sistema dinâmico é o menor conjunto de variáveis (chamadas variáveis de estado) tal que o conhecimento destas variáveis em $t = t_0$, juntamente com a entrada para $t \geq t_0$, determina completamente o comportamento do sistema para qualquer instante $t \geq t_0$.

Portanto, o estado de um sistema dinâmico no instante t é univocamente determinado pelo estado no instante t_0 e a entrada para $t \geq t_0$, e é independente do estado e a entrada antes de t_0. Note que, ao lidar com sistemas lineares invariantes no tempo, usualmente escolhemos a referência de tempo t_0 como zero.

Variáveis de estado. As variáveis de estado de um sistema dinâmico são o menor conjunto de variáveis que determina o estado do sistema dinâmico. Se pelo menos n variáveis $x_1(t), x_2(t), \ldots, x_n(t)$ são necessárias para descrever completamente o comportamento de um sistema dinâmico (tal que uma vez dada a entrada para $t \geq t_0$ e o estado inicial em $t = t_0$ é especificado, o estado futuro do sistema está completamente determinado), então as tais n variáveis $x_1(t), x_2(t), \ldots, x_n(t)$ são um conjunto de variáveis de estado. Note que as variáveis de estado não precisam ser grandezas fisicamente mensuráveis ou observáveis. Em termos práticos, entretanto, é conveniente escolher grandezas facilmente mensuráveis para as variáveis de estado porque leis de controle ótimo irão requerer a realimentação de todas as variáveis de estado com ganhos adequados.

Vetor de estado. Se n variáveis de estado são necessárias para descrever completamente o comportamento de um dado sistema, então estas n variáveis de estado podem ser consideradas como as n componentes de um vetor $\mathbf{x}(t)$. Tal vetor é chamado de um vetor de estados. Um vetor de estado é portanto um vetor que determina unicamente o estado do sistema $\mathbf{x}(t)$ para qualquer $t \geq t_0$, uma vez que a entrada $\mathbf{u}(t)$ para $t \geq t_0$ é especificada.

Espaço de estados. O espaço n-dimensional cujos eixos de coordenadas são os eixos $x_1, x_2, \ldots, x_n$ é chamado de um espaço de estados. Qualquer estado pode ser representado por um ponto no espaço de estados.

Exemplo 14.1 Considere o circuito *RLC* visto na Fig. 14.1. O comportamento dinâmico do sistema é completamente definido para $t \geq t_0$ se os valores iniciais da corrente $i(t_0)$, a tensão do capacitor $v_c(t_0)$, e a tensão de entrada $v(t)$ para $t \geq t_0$ são conhecidos. Portanto, o estado do circuito para $t \geq t_0$ é completamente determinado por $i(t)$, $v_c(t)$, e a tensão de entrada $v(t)$ para $t \geq t_0$. Portanto $i(t)$ e $v_c(t)$ são um conjunto de variáveis de estado para este sistema.

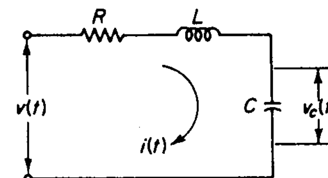


Fig. 14.1 Circuito *RLC*.

(Note, entretanto, que a escolha de variáveis de estado para um dado sistema não é única. Por exemplo, neste sistema $x_1(t) = v_c(t) + Ri(t)$ e $x_2(t) = v_c(t)$ podem ser escolhidos como um conjunto de variáveis de estado.)

Suponha que escolhemos $i(t)$ e $v_c(t)$ como as variáveis de estado. Então as equações que descrevem a dinâmica do sistema são

$$L\frac{di}{dt} + Ri + v_c = v$$

$$C\frac{dv_c}{dt} = i$$

Em notação vetorial-matricial, temos

$$\begin{bmatrix} \dot{i} \\ \dot{v}_c \end{bmatrix} = \begin{bmatrix} -\frac{R}{L} & -\frac{1}{L} \\ \frac{1}{C} & 0 \end{bmatrix} \begin{bmatrix} i \\ v_c \end{bmatrix} + \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix} [v]$$

Esta é uma representação de espaço de estados para o dado sistema ou circuito.

Análise de sistemas complexos. Um sistema moderno complexo pode ter muitas entradas e muitas saídas, e estas podem estar inter-relacionadas de uma maneira

complicada. Para analisar tal sistema, é essencial reduzir a complexidade nas expressões matemáticas, bem como utilizar computadores para a maioria dos cálculos maçantes necessários para a análise. A abordagem de espaço de estados para a análise de sistemas é mais adequada sob este ponto de vista.

Enquanto a teoria de controle convencional é baseada na relação entre entrada e saída ou função de transferência, a teoria de controle moderno se baseia na descrição das equações do sistema em termos de n equações diferenciais de primeira-ordem, que podem ser combinadas em uma equação diferencial vetorial-matricial de primeira-ordem. O uso da notação vetor-matriz simplifica muito a representação matemática de sistemas de equações. O aumento no número de variáveis de estado, de entradas, ou de saídas não aumenta a complexidade das equações. De fato, a análise de sistemas complicados de múltiplas entradas e saídas pode ser feita por procedimentos que são apenas um pouco mais complicados do que aqueles requeridos para a análise de sistemas de equações diferenciais escalares de primeira-ordem.

Do ponto de vista computacional, os métodos de espaço de estados são particularmente adequados para computações em computadores digitais por causa da abordagem no domínio do tempo. Isto libera o engenheiro de computações maçantes que seriam de outra forma necessárias e permite que ele dedique seus esforços apenas para os aspectos analíticos do problema. Esta é uma das vantagens dos métodos de espaço de estados.

Finalmente, é importante notar que não é necessário que as variáveis de estado representem grandezas físicas do sistema. Podem-se escolher como variáveis de estado grandezas que não representam grandezas físicas bem como aquelas que não são nem mensuráveis nem observáveis. Tal liberdade de escolha de variáveis de estado é outra vantagem de métodos de espaço de estados.

14.2 REPRESENTAÇÃO DE SISTEMAS POR ESPAÇO DE ESTADOS

Um sistema dinâmico que consiste em um número finito de elementos concentrados pode ser descrito por equações diferenciais ordinárias em que o tempo é a variável independente. Usando-se notação matricial, uma equação diferencial de ordem n pode ser representada por uma equação matricial diferencial de primeira-ordem. Se n elementos do vetor são um conjunto de variáveis de estado, então a equação matricial diferencial é chamada uma equação de *estado*. Nesta seção apresentaremos métodos para obtenção de representações de espaço de estados de sistemas contínuos.

Representação de espaço de estados de sistemas de equações diferenciais lineares de ordem n em que a função excitação não envolve termos em derivadas. Considere o seguinte sistema de ordem n:

$$\overset{(n)}{y} + a_1 \overset{(n-1)}{y} + \cdots + a_{n-1}\dot{y} + a_n y = u \qquad (14.1)$$

Observando que o conhecimento de $y(0), \dot{y}(0), \ldots, \overset{(n-1)}{y}(0)$, junto com a entrada $u(t)$ para $t \geq 0$, determina completamente o futuro comportamento do sistema, podemos considerar $y(t), \dot{y}(t), \ldots, \overset{(n-1)}{y}(t)$ como um conjunto de n variáveis de estado. (Matematicamente, tal escolha é bastante conveniente. Na prática, entretanto, em virtude da imprecisão de termos em derivada de ordem superior devido a efeitos de ruídos inerentes em quaisquer situações práticas, tal escolha de variáveis de estado pode não ser desejável.)

Vamos definir

$$\begin{aligned} x_1 &= y \\ x_2 &= \dot{y} \\ &\cdots \\ x_n &= \overset{(n-1)}{y} \end{aligned}$$

Então a Eq. (14.1) pode ser escrita como

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\cdots \\ \dot{x}_{n-1} &= x_n \\ \dot{x}_n &= -a_n x_1 - \cdots - a_1 x_n + u \end{aligned}$$

ou

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} - \mathbf{B}u \qquad (14.2)$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix}$$

A equação de saída se torna

$$y = [1 \quad 0 \quad \cdots \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

ou

$$y = \mathbf{C}\mathbf{x} \qquad (14.3)$$

onde

$$\mathbf{C} = [1 \quad 0 \quad \cdots \quad 0]$$

A equação diferencial de primeira-ordem, a Eq. (14.2), é a equação de estado, e a equação algébrica, a Eq. (14.3), é a equação de saída.

Exemplo 14.2 Considere o sistema definido por

$$\dddot{y} + 6\ddot{y} + 11\dot{y} + 6y = 6u \tag{14.4}$$

onde y é a saída e u é a entrada do sistema. Obtenha uma representação em espaço de estados do sistema.

Vamos escolher as variáveis de estado como

$$x_1 = y$$
$$x_2 = \dot{y}$$
$$x_3 = \ddot{y}$$

Então obtemos

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = x_3$$
$$\dot{x}_3 = -6x_1 - 11x_2 - 6x_3 + 6u$$

A última destas três equações foi obtida resolvendo-se a equação diferencial original para o termo em derivada de ordem mais alta $\dddot{y}$ e então substituindo $y = x_1, \dot{y} = x_2, \ddot{y} = x_3$ na equação resultante. Usando-se a notação matricial, estas três equações diferenciais de primeira-ordem podem ser combinadas em uma só como segue:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix} [u] \tag{14.5}$$

A equação de saída é dada por

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{14.6}$$

As Eqs. (14.5) e (14.6) podem ser postas em uma forma padrão como

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \tag{14.7}$$

$$y = \mathbf{C}\mathbf{x} \tag{14.8}$$

onde

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix}, \quad \mathbf{C} = [1 \quad 0 \quad 0]$$

A Fig. 14.2 mostra a representação em diagrama de blocos da equação de estados e da

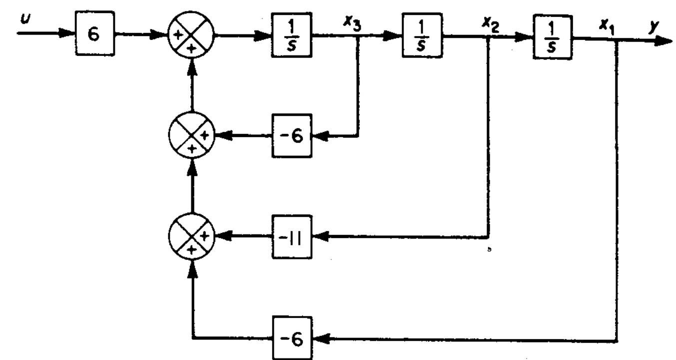


Fig. 14.2 Representação em diagrama em blocos do sistema descrito pelas Eqs. (14.7) e (14.8).

equação de saída. Observe que as funções de transferência dos blocos de realimentação são idênticas aos valores negativos dos coeficientes da equação diferencial original, a Eq. (14.4).

A não-unicidade do conjunto de variáveis de estado. Dissemos que um conjunto de variáveis de estado não é único para um dado sistema. Suponha que $x_1, x_2, \ldots, x_n$ são um conjunto de variáveis de estado. Então podemos considerar como outro conjunto de variáveis de estado qualquer conjunto de funções.

$$\hat{x}_1 = X_1(x_1, x_2, \ldots, x_n)$$
$$\hat{x}_2 = X_2(x_1, x_2, \ldots, x_n)$$
$$\cdots$$
$$\hat{x}_n = X_n(x_1, x_2, \ldots, x_n)$$

contanto que, para cada conjunto de valores $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n$, corresponda um único conjunto de valores $x_1, x_2, \ldots, x_n$ e vice-versa. Portanto, se $\mathbf{x}$ é um vetor de estado, então $\hat{\mathbf{x}}$, onde

$$\hat{\mathbf{x}} = \mathbf{P}\mathbf{x}$$

também é um vetor de estados, contanto que $\mathbf{P}$ seja não-singular. Diferentes vetores de estado fornecem a mesma informação sobre o comportamento do sistema.

Autovalores de uma matriz $\mathbf{A}$ $n \times n$. Os autovalores de uma matriz $\mathbf{A}$ $n \times n$ são as raízes da equação característica

$$|\lambda\mathbf{I} - \mathbf{A}| = 0$$

Os autovalores são às vezes chamados de raízes características.

Considere, por exemplo, a seguinte matriz $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}$$

A equação característica é

$$|\lambda\mathbf{I} - \mathbf{A}| = \begin{vmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ 6 & 11 & \lambda + 6 \end{vmatrix}$$
$$= \lambda^3 + 6\lambda^2 + 11\lambda + 6$$
$$= (\lambda + 1)(\lambda + 2)(\lambda + 3) = 0$$

Os autovalores de **A** são as raízes da equação característica, ou -1, -2, e -3.

Exemplo 14.3 Considere o mesmo sistema discutido no Exemplo 14.2. Mostraremos que a Eq. (14.5) não é a única equação de estado possível para este sistema. Suponha que definamos um novo conjunto de variáveis de estado z_1, z_2, z_3 pela transformação

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & -3 \\ 1 & 4 & 9 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

ou

$$\mathbf{x} = \mathbf{P}\mathbf{z} \tag{14.9}$$

onde

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & -3 \\ 1 & 4 & 9 \end{bmatrix} \tag{14.10}$$

Então, substituindo a Eq. (14.9) na Eq. (14.7), obtemos

$$\mathbf{P}\dot{\mathbf{z}} = \mathbf{A}\mathbf{P}\mathbf{z} + \mathbf{B}u$$

Multiplicando ambos os lados desta última equação à esquerda por $\mathbf{P}^{-1}$, obtemos

$$\dot{\mathbf{z}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{z} + \mathbf{P}^{-1}\mathbf{B}u \tag{14.11}$$

ou

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} 3 & 2{,}5 & 0{,}5 \\ -3 & -4 & -1 \\ 1 & 1{,}5 & 0{,}5 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & -3 \\ 1 & 4 & 9 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$
$$+ \begin{bmatrix} 3 & 2{,}5 & 0{,}5 \\ -3 & -4 & -1 \\ 1 & 1{,}5 & 0{,}5 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix} [u]$$

Simplificando, temos

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix} [u] \tag{14.12}$$

A Eq. (14.12) também é uma equação de estados que descreve o mesmo sistema definido pela Eq. (14.5).

A equação de saída, a Eq. (14.8), é modificada para

$$y = \mathbf{C}\mathbf{P}\mathbf{z}$$

ou

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & -3 \\ 1 & 4 & 9 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$
$$= [1 \quad 1 \quad 1] \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \tag{14.13}$$

Observe que a matriz de transformação **P**, definida pela Eq. (14.10), transforma a matriz de coeficientes de **z** na matriz diagonal. Como claramente visto na Eq. (14.12), as três equações de estado separadas são desacopladas. Observe também que os elementos diagonais da matriz $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ na Eq. (14.11) são idênticos aos três autovalores de **A**. É muito importante notar que os autovalores de **A** e aqueles de $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ são idênticos. A seguir provaremos isto para um caso geral.

Invariância dos autovalores. Para provar a invariância dos autovalores sob uma transformação linear, devemos mostrar que os polinômios característicos $|\lambda\mathbf{I} - \mathbf{A}|$ e $|\lambda\mathbf{I} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P}|$ são idênticos.

Como o determinante de um produto é o produto dos determinantes, obtemos

$$|\lambda\mathbf{I} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P}| = |\lambda\mathbf{P}^{-1}\mathbf{P} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P}|$$
$$= |\mathbf{P}^{-1}(\lambda\mathbf{I} - \mathbf{A})\mathbf{P}|$$
$$= |\mathbf{P}^{-1}||\lambda\mathbf{I} - \mathbf{A}||\mathbf{P}|$$
$$= |\mathbf{P}^{-1}||\mathbf{P}||\lambda\mathbf{I} - \mathbf{A}|$$

Observando que o produto dos determinantes $|\mathbf{P}^{-1}|$ e $|\mathbf{P}|$ é o determinante do produto $|\mathbf{P}^{-1}\mathbf{P}|$, obtemos

$$|\lambda\mathbf{I} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P}| = |\mathbf{P}^{-1}\mathbf{P}||\lambda\mathbf{I} - \mathbf{A}|$$
$$= |\lambda\mathbf{I} - \mathbf{A}|$$

Portanto provamos que os autovalores de **A** são invariantes sob uma transformação linear.

Diagonalização de matrizes $n \times n$. Note que se uma matriz **A** $n \times n$ com autovalores distintos é dada por

$$
\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \tag{14.14}
$$

a transformação $\mathbf{x} = \mathbf{Pz}$ onde

$$
\mathbf{P} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix}
$$

$\lambda_1, \lambda_2, \ldots, \lambda_n = n$ autovalores distintos de **A**

transformará $\mathbf{P}^{-1}\mathbf{AP}$ em uma matriz diagonal, ou

$$
\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} \lambda_1 & & & & 0 \\ & \lambda_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ 0 & & & & \lambda_n \end{bmatrix}
$$

Se a matriz **A** definida pela Eq. (14.14) envolve autovalores múltiplos então a diagonalização é impossível. Por exemplo, se a matriz **A** 3×3, onde

$$
\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{bmatrix}
$$

tem os autovalores $\lambda_1, \lambda_1, \lambda_3$, então a transformação $\mathbf{x} = \mathbf{Sz}$, onde

$$
\mathbf{S} = \begin{bmatrix} 1 & 0 & 1 \\ \lambda_1 & 1 & \lambda_3 \\ \lambda_1^2 & 2\lambda_1 & \lambda_3^2 \end{bmatrix}
$$

fornecerá

$$
\mathbf{S}^{-1}\mathbf{AS} = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}
$$

Esta é a chamada forma de Jordan canônica.

Exemplo 14.4 Considere o mesmo sistema discutido nos Exemplos 14.2 e 14.3, reescrito abaixo:

$$
\dddot{y} + 6\ddot{y} + 11\dot{y} + 6y = 6u \tag{14.15}
$$

Demonstraremos que a representação de espaço de estado dada pelas Eqs. (14.12) e (14.13) pode também ser obtida usando-se a técnica da expansão em frações parciais.

Vamos reescrever a Eq. (14.15) na forma de uma função de transferência:

$$
\frac{Y(s)}{U(s)} = \frac{6}{s^3 + 6s^2 + 11s + 6} = \frac{6}{(s+1)(s+2)(s+3)}
$$

Expandindo esta função de transferência em frações parciais, obtemos

$$
\frac{Y(s)}{U(s)} = \frac{3}{s+1} + \frac{-6}{s+2} + \frac{3}{s+3}
$$

Portanto

$$
Y(s) = \frac{3}{s+1}U(s) + \frac{-6}{s+2}U(s) + \frac{3}{s+3}U(s) \tag{14.16}
$$

Vamos definir

$$
X_1(s) = \frac{3}{s+1}U(s) \tag{14.17}
$$

$$
X_2(s) = \frac{-6}{s+2}U(s) \tag{14.18}
$$

$$
X_3(s) = \frac{3}{s+3}U(s) \tag{14.19}
$$

As transformadas inversas de Laplace das Eqs. (14.17), (14.18) e (14.19) fornecem

$$
\dot{x}_1 = -x_1 + 3u
$$

$$
\dot{x}_2 = -2x_2 - 6u
$$

$$
\dot{x}_3 = -3x_3 + 3u
$$

Como a Eq. (14.16) pode ser escrita como

$$
Y(s) = X_1(s) + X_2(s) + X_3(s)
$$

obtemos

$$y = x_1 + x_2 + x_3$$

Em termos da notação matricial, obtemos

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix} [u] \tag{14.20}$$

$$y = [1 \quad 1 \quad 1] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{14.21}$$

As Eqs. (14.20) e (14.21) são claramente idênticas às Eqs. (14.12) e (14.13), respectivamente.

A Fig. 14.3 mostra uma representação em diagramas de blocos das Eqs. (14.20) e (14.21). Note que as funções de transferência nos blocos de realimentação são idênticas aos autovalores do sistema. Note também que os resíduos dos pólos da função de transferência, ou os coeficientes da expansão em frações parciais de $Y(s)/U(s)$, aparecem nos blocos de elo direto.

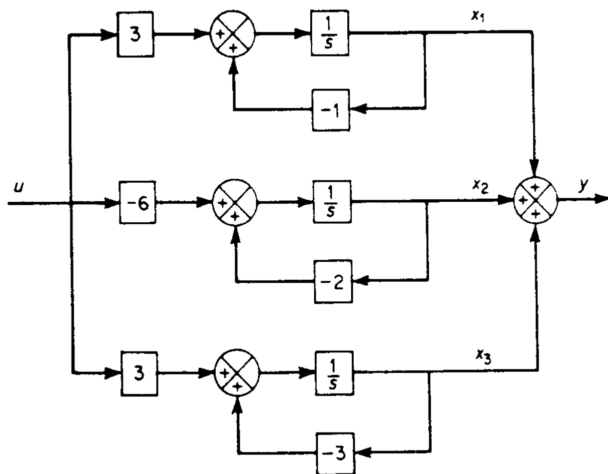


Fig. 14.3 Representação em diagrama em blocos do sistema descrito pelas Eqs. (14.20) e (14.21).

Representação de espaço de estados de sistemas de equações diferenciais de ordem n com r funções de excitação. Considere o sistema de múltiplas entradas e saídas visto na Fig. 14.4. Neste sistema, $x_1, x_2, \ldots, x_n$ representam as variáveis de estado; $u_1, u_2, \ldots, u_r$ representam as variáveis de entrada; e $y_1, y_2, \ldots, y_m$ são as variáveis de saída. Da Fig. 14.4, obtemos o sistema de equações como segue:

$$\begin{aligned} \dot{x}_1 &= a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n \\ &\quad + b_{11}(t)u_1 + b_{12}(t)u_2 + \cdots + b_{1r}(t)u_r \\ \dot{x}_2 &= a_{21}(t)x_1 + a_{22}(t)x_2 + \cdots + a_{2n}(t)x_n \\ &\quad + b_{21}(t)u_1 + b_{22}(t)u_2 + \cdots + b_{2r}(t)u_r \\ &\cdots \\ \dot{x}_n &= a_{n1}(t)x_1 + a_{n2}(t)x_2 + \cdots + a_{nn}(t)x_n \\ &\quad + b_{n1}(t)u_1 + b_{n2}(t)u_2 + \cdots + b_{nr}(t)u_r \end{aligned}$$

onde os $a(t)$ e $b(t)$ são constantes ou funções de t. Em termos da notação matricial, estas n equações podem ser escritas compactamente como

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u} \tag{14.22}$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \text{vetor de estado}$$

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_r \end{bmatrix} = \text{vetor de entrada (ou controle)}$$

$$\mathbf{A}(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{bmatrix}$$

$$\mathbf{B}(t) = \begin{bmatrix} b_{11}(t) & b_{12}(t) & \cdots & b_{1r}(t) \\ b_{21}(t) & b_{22}(t) & \cdots & b_{2r}(t) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ b_{n1}(t) & b_{n2}(t) & \cdots & b_{nr}(t) \end{bmatrix}$$

A Eq. (14.22) é a equação de estado para o sistema. [Note que uma equação diferencial matricial, como a da Eq. (14.22) (ou as n equações diferenciais de

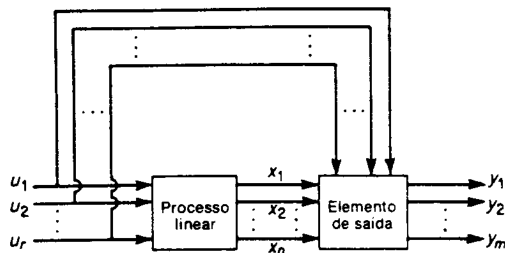


Fig. 14.4 Sistema de múltiplas entradas e múltiplas saídas.

primeira-ordem equivalentes), descrevendo a dinâmica de um sistema, é uma equação de estado se e apenas se o conjunto de variáveis dependentes na equação diferencial matricial satisfaz a definição de variáveis de estado.]

Para os sinais de saída, obtemos

$$y_1 = c_{11}(t)x_1 + c_{12}(t)x_2 + \cdots + c_{1n}(t)x_n + d_{11}(t)u_1 + d_{12}(t)u_2 + \cdots + d_{1r}(t)u_r$$

$$y_2 = c_{21}(t)x_1 + c_{22}(t)x_2 + \cdots + c_{2n}(t)x_n + d_{21}(t)u_1 + d_{22}(t)u_2 + \cdots + d_{2r}(t)u_r$$

$$\cdots$$

$$y_m = c_{m1}(t)x_1 + c_{m2}(t)x_2 + \cdots + c_{mn}(t)x_n + d_{m1}(t)u_1 + d_{m2}(t)u_2 + \cdots + d_{mr}(t)u_r$$

Em termos de notação matricial, estas m equações podem ser escritas compactamente como

$$\mathbf{y} = \mathbf{C}(t)\mathbf{x} + \mathbf{D}(t)\mathbf{u} \tag{14.23}$$

onde

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix} = \text{vetor de saída}$$

$$\mathbf{C}(t) = \begin{bmatrix} c_{11}(t) & c_{12}(t) & \cdots & c_{1n}(t) \\ c_{21}(t) & c_{22}(t) & \cdots & c_{2n}(t) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ c_{m1}(t) & c_{m2}(t) & \cdots & c_{mn}(t) \end{bmatrix}$$

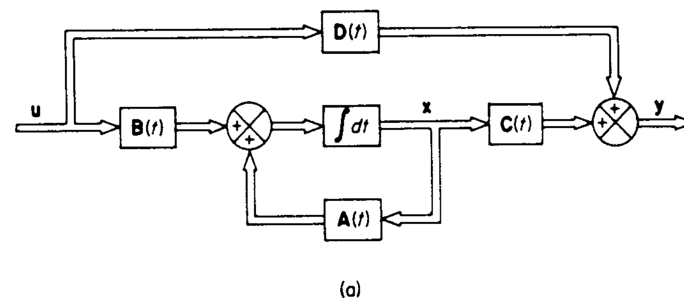
$$\mathbf{D}(t) = \begin{bmatrix} d_{11}(t) & d_{12}(t) & \cdots & d_{1r}(t) \\ d_{21}(t) & d_{22}(t) & \cdots & d_{2r}(t) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ d_{m1}(t) & d_{m2}(t) & \cdots & d_{mr}(t) \end{bmatrix}$$

a Eq. (14.23) é a equação de saída do sistema. As matrizes $\mathbf{A}(t)$, $\mathbf{B}(t)$, $\mathbf{C}(t)$ e $\mathbf{D}(t)$ caracterizam completamente a dinâmica do sistema.

Uma representação em diagrama de blocos e uma representação em gráfico de fluxo de sinais do sistema definido pelas Eqs. (14.22) e (14.23) são vistas nas Figs. 14.5(a) e (b), respectivamente. Para indicar grandezas vetoriais, usamos duplas flechas nos diagramas.

Representação de espaço de estados de sistemas de equações diferenciais de ordem n em que a função excitação envolve termos em derivada. Se a equação diferencial do sistema envolve derivadas da função de excitação, como por exemplo.

$$\overset{(n)}{y} + a_1\overset{(n-1)}{y} + \cdots + a_{n-1}\dot{y} + a_n y = b_0\overset{(n)}{u} + b_1\overset{(n-1)}{u} + \cdots + b_{n-1}\dot{u} + b_n u \tag{14.24}$$



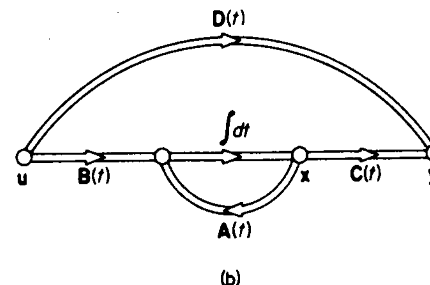


Fig. 14.5 (a) Representação em diagrama em blocos do sistema descrito pelas Eqs. (14.22) e (14.23); (b) representação por diagrama de fluxo de sinal do sistema de (a).

então o conjunto de n variáveis $y, \dot{y}, \ddot{y}, \ldots, \overset{(n-1)}{y}$ não é um conjunto de variáveis de estado. e o método simples que empregamos anteriormente não pode ser empregado. Isto porque n equações diferenciais de primeira-ordem

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_3 \\
&\cdots \\
\dot{x}_n &= -a_n x_1 - a_{n-1} x_2 - \cdots - a_1 x_n + b_0 \overset{(n)}{u} + b_1 \overset{(n-1)}{u} + \cdots + b_n u
\end{aligned}$$

onde $x_1 = y$ podem não fornecer uma solução única.

O problema principal em se definir as variáveis de estado para este caso está na existência de termos em derivada no lado direito da última equação do sistema de n equações acima. As variáveis de estado devem ser tais que eliminem as derivadas de u na equação de estado.

Um fato bem conhecido na teoria de controle moderno é que se definirmos as seguintes n variáveis como um conjunto de n variáveis de estado.

$$\begin{aligned}
x_1 &= y - \beta_0 u \\
x_2 &= \dot{y} - \beta_0 \dot{u} - \beta_1 u = \dot{x}_1 - \beta_1 u \\
x_3 &= \ddot{y} - \beta_0 \ddot{u} - \beta_1 \dot{u} - \beta_2 u = \dot{x}_2 - \beta_2 u \\
&\cdots \\
x_n &= \overset{(n-1)}{y} - \beta_0 \overset{(n-1)}{u} - \beta_1 \overset{(n-2)}{u} - \cdots - \beta_{n-2} \dot{u} - \beta_{n-1} u = \dot{x}_{n-1} - \beta_{n-1} u
\end{aligned} \tag{14.25}$$

onde $\beta_0, \beta_1, \ldots, \beta_n$ são determinadas de

$$\begin{aligned}
\beta_0 &= b_0 \\
\beta_1 &= b_1 - a_1 \beta_0 \\
\beta_2 &= b_2 - a_1 \beta_1 - a_2 \beta_0 \\
\beta_3 &= b_3 - a_1 \beta_2 - a_2 \beta_1 - a_3 \beta_0 \\
&\cdots \\
\beta_n &= b_n - a_1 \beta_{n-1} - \cdots - a_{n-1} \beta_1 - a_n \beta_0
\end{aligned} \tag{14.26}$$

então a existência e unicidade da solução da equação de estado estão asseguradas. (Note que esta não é a única escolha de um conjunto de variáveis de estado.) Com a presente escolha de variáveis de estado, obtemos a seguinte equação de estado e equação de saída para o sistema da Eq. (14.24):

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_{n-1} \\ \beta_n \end{bmatrix} [u]$$

$$y = [1 \quad 0 \quad \cdots \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} + \beta_0 u$$

ou

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \tag{14.27}$$

$$y = \mathbf{C}\mathbf{x} + Du \tag{14.28}$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1} \\ x_n \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_{n-1} \\ \beta_n \end{bmatrix}, \qquad \mathbf{C} = [1 \quad 0 \quad \cdots \quad 0], \qquad D = \beta_0 = b_0$$

A condição inicial $\mathbf{x}(0)$ pode ser determinada usando-se a Eq. (14.25).

Nesta representação de espaço de estados. a matriz **A** é essencialmente a mesma que no sistema da Eq. (14.1). As derivadas no lado direito da Eq. (14.24) afetam apenas os elementos da matriz **B**.

Note que a representação de espaço de estados para a seguinte função de transferência

$$\frac{Y(s)}{U(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \cdots + b_{n-1} s + b_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n}$$

é dada também pelas Eqs. (14.27) e (14.28).

Exemplo 14.5 Considere o sistema de controle visto na Fig. 14.6. A função de transferência de malha-fechada é

$$\frac{Y(s)}{U(s)} = \frac{160(s+4)}{s^3 + 18s^2 + 192s + 640}$$

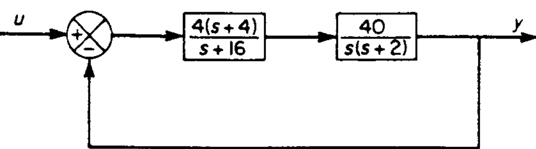


Fig. 14.6 Sistema de controle.

A correspondente equação diferencial é

$$\dddot{y} + 18\ddot{y} + 192\dot{y} + 640y = 160\dot{u} + 640u$$

Obtenha uma representação de espaço de estado do sistema.

Referindo-se à Eq. (14.25), vamos definir

$$\begin{aligned} x_1 &= y - \beta_0 u \\ x_2 &= \dot{y} - \beta_0 \dot{u} - \beta_1 u = \dot{x}_1 - \beta_1 u \\ x_3 &= \ddot{y} - \beta_0 \ddot{u} - \beta_1 \dot{u} - \beta_2 u = \dot{x}_2 - \beta_2 u \end{aligned}$$

onde β_0, β_1 e β_2 são determinados da Eq. (14.26) como segue:

$$\begin{aligned} \beta_0 &= b_0 = 0 \\ \beta_1 &= b_1 - a_1\beta_0 = 0 \\ \beta_2 &= b_2 - a_1\beta_1 - a_2\beta_0 = 160 \\ \beta_3 &= b_3 - a_1\beta_2 - a_2\beta_1 - a_3\beta_0 = -2240 \end{aligned}$$

Então a equação de estado para o sistema se torna

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -640 & -192 & -18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 160 \\ -2240 \end{bmatrix} [u]$$

A equação de saída se torna

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

14.3 SOLUÇÃO DA EQUAÇÃO DE ESTADO INVARIANTE NO TEMPO

Nesta seção, obteremos a solução geral da equação de estado linear invariante no tempo. Consideraremos inicialmente o caso homogêneo, e então o caso não homogêneo.

Solução de equações de estado homogêneas. Antes de resolver equações diferenciais matriciais, vamos rever a solução da equação diferencial escalar

$$\dot{x} = ax \tag{14.29}$$

Ao resolver esta equação, podemos supor uma solução $x(t)$ da forma

$$x(t) = b_0 + b_1 t + b_2 t^2 + \cdots + b_k t^k + \cdots \tag{14.30}$$

Substituindo esta solução suposta na Eq. (14.29), obtemos

$$\begin{aligned} &b_1 + 2b_2 t + 3b_3 t^2 + \cdots + k b_k t^{k-1} + \cdots \\ &= a(b_0 + b_1 t + b_2 t^2 + \cdots + b_k t^k + \cdots) \end{aligned} \tag{14.31}$$

Se a solução suposta deve ser a solução verdadeira, a Eq. (14.31) deve valer para qualquer t. Portanto, equacionando os coeficientes de iguais potências de t, obtemos

$$\begin{aligned} b_1 &= ab_0 \\ b_2 &= \frac{1}{2}ab_1 = \frac{1}{2}a^2 b_0 \\ b_3 &= \frac{1}{3}ab_2 = \frac{1}{3 \times 2}a^3 b_0 \\ &\cdots \\ b_k &= \frac{1}{k!}a^k b_0 \end{aligned}$$

O valor de b_0 é determinado substituindo-se $t = 0$ na Eq. (14.30), ou

$$x(0) = b_0$$

Portanto a solução $x(t)$ pode ser escrita como

$$x(t) = \left(1 + at + \frac{1}{2!}a^2t^2 + \cdots + \frac{1}{k!}a^k t^k + \cdots\right)x(0) = e^{at}x(0)$$

Resolveremos agora a equação diferencial matricial

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{14.32}$$

onde

$\mathbf{x}$ = vetor n-dimensional

$\mathbf{A}$ = matriz constante $n \times n$

Por analogia com o caso escalar, suporemos que a solução é na forma de uma série de potências vetoriais em t, ou

$$\mathbf{x}(t) = \mathbf{b}_0 + \mathbf{b}_1 t + \mathbf{b}_2 t^2 + \cdots + \mathbf{b}_k t^k + \cdots \tag{14.33}$$

Substituindo esta suposta solução na Eq. (14.32), obtemos

$$\begin{aligned} &\mathbf{b}_1 + 2\mathbf{b}_2 t + 3\mathbf{b}_3 t^2 + \cdots + k\mathbf{b}_k t^{k-1} + \cdots \\ &\qquad = \mathbf{A}(\mathbf{b}_0 + \mathbf{b}_1 t + \mathbf{b}_2 t^2 + \cdots + k\mathbf{b}_k t^k + \cdots) \end{aligned} \tag{14.34}$$

Se a solução suposta deve ser a solução verdadeira. a Eq. (14.34) valer valor para todo t. Portanto impomos que os coeficientes de potências iguais de t sejam idênticos. ou

$$\mathbf{b}_1 = \mathbf{A}\mathbf{b}_0$$

$$\mathbf{b}_2 = \frac{1}{2}\mathbf{A}\mathbf{b}_1 = \frac{1}{2}\mathbf{A}^2\mathbf{b}_0$$

$$\mathbf{b}_3 = \frac{1}{3}\mathbf{A}\mathbf{b}_2 = \frac{1}{3 \times 2}\mathbf{A}^3\mathbf{b}_0$$

$$\cdots$$

$$\mathbf{b}_k = \frac{1}{k!}\mathbf{A}^k\mathbf{b}_0$$

Substituindo $t = 0$ na Eq. (14.33), obtemos

$$\mathbf{x}(0) = \mathbf{b}_0$$

Portanto a solução $\mathbf{x}(t)$ pode ser escrita como

$$\mathbf{x}(t) = \left(\mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2t^2 + \cdots + \frac{1}{k!}\mathbf{A}^kt^k + \cdots\right)\mathbf{x}(0)$$

A expressão em parênteses no lado direito desta última equação é uma matriz $n \times n$. Em virtude da sua similitude com a série infinita de potências para uma exponencial escalar, chamamos esta de matriz exponencial e escrevemos

$$\mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2t^2 + \cdots + \frac{1}{k!}\mathbf{A}^kt^k + \cdots = e^{\mathbf{A}t}$$

Em termos da matriz exponencial, a solução da Eq. (14.32) pode ser escrita como

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) \tag{14.35}$$

Como a matriz exponencial é muito importante na análise de espaço de estados de sistemas lineares, examinaremos em seguida as propriedades da matriz exponencial.

Matriz exponencial. Pode-se provar que a matriz exponencial de uma matriz $\mathbf{A}$ $n \times n$

$$e^{\mathbf{A}t} = \sum_{k=0}^{\infty}\frac{\mathbf{A}^kt^k}{k!}$$

converge absolutamente para todo t finito. (Portanto cálculos em computador para calcular os elementos de $e^{\mathbf{A}t}$, usando a expansão em série, podem ser facilmente feitos.)

Devido à convergência da série infinita $\sum_{k=0}^{\infty}\mathbf{A}^kt^k/k!$, a série pode ser diferenciada termo a termo, resultando

$$\frac{d}{dt}e^{\mathbf{A}t} = \mathbf{A} + \mathbf{A}^2t + \frac{\mathbf{A}^3t^2}{2!} + \cdots + \frac{\mathbf{A}^kt^{k-1}}{(k-1)!} + \cdots$$

$$= \mathbf{A}\left[\mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2t^2}{2!} + \cdots + \frac{\mathbf{A}^{k-1}t^{k-1}}{(k-1)!} + \cdots\right] = \mathbf{A}e^{\mathbf{A}t}$$

$$= \left[\mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2t^2}{2!} + \cdots + \frac{\mathbf{A}^{k-1}t^{k-1}}{(k-1)!} + \cdots\right]\mathbf{A} = e^{\mathbf{A}t}\mathbf{A}$$

A matriz exponencial tem a seguinte propriedade:

$$e^{\mathbf{A}(t+s)} = e^{\mathbf{A}t}e^{\mathbf{A}s}$$

Isto pode ser provado como segue:

$$e^{\mathbf{A}t}e^{\mathbf{A}s} = \left(\sum_{k=0}^{\infty}\frac{\mathbf{A}^kt^k}{k!}\right)\left(\sum_{k=0}^{\infty}\frac{\mathbf{A}^ks^k}{k!}\right)$$

$$= \sum_{k=0}^{\infty}\mathbf{A}^k\left(\sum_{i=0}^{k}\frac{t^is^{k-i}}{i!(k-i)!}\right)$$

$$= \sum_{k=0}^{\infty}\mathbf{A}^k\frac{(t+s)^k}{k!}$$

$$= e^{\mathbf{A}(t+s)}$$

Em particular, se $s = -t$, então

$$e^{\mathbf{A}t}e^{-\mathbf{A}t} = e^{-\mathbf{A}t}e^{\mathbf{A}t} = e^{\mathbf{A}(t-t)} = \mathbf{I}$$

Portanto a inversa de $e^{\mathbf{A}t}$ é $e^{-\mathbf{A}t}$. Como a inversa de $e^{\mathbf{A}t}$ sempre existe, $e^{\mathbf{A}t}$ é não singular.

É muito importante lembrar que

$$e^{(\mathbf{A}+\mathbf{B})t} = e^{\mathbf{A}t}e^{\mathbf{B}t} \qquad \text{se } \mathbf{AB} = \mathbf{BA}$$

$$e^{(\mathbf{A}+\mathbf{B})t} \neq e^{\mathbf{A}t}e^{\mathbf{B}t} \qquad \text{se } \mathbf{AB} \neq \mathbf{BA}$$

Para provar isto, notamos que

$$e^{(\mathbf{A}+\mathbf{B})t} = \mathbf{I} + (\mathbf{A}+\mathbf{B})t + \frac{(\mathbf{A}+\mathbf{B})^2}{2!}t^2 + \frac{(\mathbf{A}+\mathbf{B})^3}{3!}t^3 + \cdots$$

$$e^{\mathbf{A}t}e^{\mathbf{B}t} = \left(\mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2t^2}{2!} + \frac{\mathbf{A}^3t^3}{3!} + \cdots\right)\left(\mathbf{I} + \mathbf{B}t + \frac{\mathbf{B}^2t^2}{2!} + \frac{\mathbf{B}^3t^3}{3!} + \cdots\right)$$

$$= \mathbf{I} + (\mathbf{A}+\mathbf{B})t + \frac{\mathbf{A}^2t^2}{2!} + \mathbf{AB}t^2 + \frac{\mathbf{B}^2t^2}{2!} + \frac{\mathbf{A}^3t^3}{3!}$$

$$+ \frac{\mathbf{A}^2\mathbf{B}t^3}{2!} + \frac{\mathbf{AB}^2t^3}{2!} + \frac{\mathbf{B}^3t^3}{3!} + \cdots$$

Portanto

$$e^{(\mathbf{A}+\mathbf{B})t} - e^{\mathbf{A}t}e^{\mathbf{B}t} = \frac{\mathbf{BA} - \mathbf{AB}}{2!}t^2 + \frac{\mathbf{BA}^2 + \mathbf{ABA} - \mathbf{B}^2\mathbf{A} + \mathbf{BAB} - 2\mathbf{A}^2\mathbf{B} - 2\mathbf{AB}^2}{3!}t^3 + \cdots$$

A diferença entre $e^{(\mathbf{A}+\mathbf{B})t}$ e $e^{\mathbf{A}t}e^{\mathbf{B}t}$ é nula se **A** e **B** comutam.

Abordagem da transformada de Laplace para a solução de equações de estado homogêneas. Vamos inicialmente considerar o caso escalar:

$$\dot{x} = ax \tag{14.36}$$

Aplicando a transformada de Laplace para a Eq. (14.36), obtemos

$$sX(s) - x(0) = aX(s) \tag{14.37}$$

onde $X(s) = \mathscr{L}[x]$. Resolvendo a Eq. (14.37) para $X(s)$, resulta

$$X(s) = \frac{x(0)}{s-a} = (s-a)^{-1}x(0)$$

A transformada inversa de Laplace desta última equação fornece a solução

$$x(t) = e^{at}x(0)$$

Esta abordagem para a solução da equação diferencial homogênea escalar pode ser estendida para a equação de estado homogênea:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) \tag{14.38}$$

Aplicando a transformada de Laplace a ambos os lados da Eq. (14.38), resulta

$$s\mathbf{X}(s) - \mathbf{x}(0) = \mathbf{A}\mathbf{X}(s)$$

onde $\mathbf{X}(s) = \mathscr{L}[\mathbf{x}]$. Portanto,

$$(s\mathbf{I} - \mathbf{A})\mathbf{X}(s) = \mathbf{x}(0)$$

Pré-multiplicando ambos os lados desta última equação por $(s\mathbf{I} - \mathbf{A})^{-1}$, resulta

$$\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0)$$

A transformada inversa de Laplace de $\mathbf{X}(s)$ fornece a solução $\mathbf{x}(t)$. Portanto,

$$\mathbf{x}(t) = \mathscr{L}^{-1}[(s\mathbf{I} - \mathbf{A})^{-1}]\mathbf{x}(0) \tag{14.39}$$

Note que

$$(s\mathbf{I} - \mathbf{A})^{-1} = \frac{\mathbf{I}}{s} + \frac{\mathbf{A}}{s^2} + \frac{\mathbf{A}^2}{s^3} + \cdots$$

Portanto, a transformada inversa de Laplace de $(s\mathbf{I} - \mathbf{A})^{-1}$ fornece

$$\mathscr{L}^{-1}[(s\mathbf{I} - \mathbf{A})^{-1}] = \mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2t^2}{2!} + \frac{\mathbf{A}^3t^3}{3!} + \cdots = e^{\mathbf{A}t} \tag{14.40}$$

(A transformada inversa de Laplace de uma matriz é a matriz que consiste nas transformadas inversas de Laplace de todos os elementos.) Das Eqs. (14.39) e (14.40), a solução da Eq. (14.38) é obtida como

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0)$$

A importância da Eq. (14.40) está no fato de que ela provê um meio conveniente de se achar a solução em forma fechada para a matriz exponencial.

Matriz de transição de estados. Podemos escrever a solução da equação de estado homogênea como

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{14.41}$$

como

$$\mathbf{x}(t) = \mathbf{\Phi}(t)\mathbf{x}(0) \tag{14.42}$$

onde $\mathbf{\Phi}(t)$ é uma matriz $n \times n$ e a solução única de

$$\dot{\mathbf{\Phi}}(t) = \mathbf{A}\mathbf{\Phi}(t), \qquad \mathbf{\Phi}(0) = \mathbf{I}$$

Para verificar isto, observe que

$$\mathbf{x}(0) = \mathbf{\Phi}(0)\mathbf{x}(0) = \mathbf{I}\mathbf{x}(0)$$

e

$$\dot{\mathbf{x}}(t) = \dot{\mathbf{\Phi}}(t)\mathbf{x}(0) = \mathbf{A}\mathbf{\Phi}(t)\mathbf{x}(0) = \mathbf{A}\mathbf{x}(t)$$

Portanto confirmamos que a Eq. (14.42) é a solução da Eq. (14.41).

Das Eqs. (14.35), (14.39) e (14.42), obtemos

$$\mathbf{\Phi}(t) = e^{\mathbf{A}t} = \mathscr{L}^{-1}[(s\mathbf{I} - \mathbf{A})^{-1}]$$

Note que

$$\mathbf{\Phi}^{-1}(t) = e^{-\mathbf{A}t} = \mathbf{\Phi}(-t)$$

Da Eq. (14.42), vemos que a solução da Eq. (14.41) é simplesmente uma transformação da condição inicial. Portanto a matriz única $\mathbf{\Phi}(t)$ é chamada de matriz de transição de estados. A matriz de transição de estados contém toda a informação sobre movimentos livres do sistema definido pela Eq. (14.41).

Se os autovalores $\lambda_1, \lambda_2, \ldots, \lambda_n$ da matriz **A** são distintos, então $\mathbf{\Phi}(t)$ conterá as n exponenciais

$$e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots, e^{\lambda_n t}$$

Em particular, se a matriz **A** é diagonal, então

$$\mathbf{\Phi}(t) = e^{\mathbf{A}t} = \begin{bmatrix} e^{\lambda_1 t} & & & & 0 \\ & e^{\lambda_2 t} & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & e^{\lambda_n t} \end{bmatrix} \quad (\mathbf{A}\text{: diagonal})$$

Se houver multiplicidade nos autovalores, por exemplo, se os autovalores de **A** são

$$\lambda_1, \lambda_1, \lambda_1, \lambda_4, \lambda_5, \ldots, \lambda_n$$

então $\mathbf{\Phi}(t)$ conterá, além das exponenciais $e^{\lambda_1 t}$, $e^{\lambda_4 t}$, $e^{\lambda_5 t}$, ..., $e^{\lambda_n t}$, termos como $te^{\lambda_1 t}$ e $t^2 e^{\lambda_1 t}$.

Propriedades de matrizes de transição de estado. Resumiremos agora as propriedades importantes da matriz de transição de estados $\mathbf{\Phi}(t)$. Para o sistema invariante no tempo

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$$

para o qual

$$\mathbf{\Phi}(t) = e^{\mathbf{A}t}$$

temos

1. $\mathbf{\Phi}(0) = e^{\mathbf{A}0} = \mathbf{I}$
2. $\mathbf{\Phi}(t) = e^{\mathbf{A}t} = (e^{-\mathbf{A}t})^{-1} = [\mathbf{\Phi}(-t)]^{-1}$

   ou $\mathbf{\Phi}^{-1}(t) = \mathbf{\Phi}(-t)$
3. $\mathbf{\Phi}(t_1 + t_2) = e^{\mathbf{A}(t_1+t_2)} = e^{\mathbf{A}t_1}e^{\mathbf{A}t_2} = \mathbf{\Phi}(t_1)\mathbf{\Phi}(t_2) = \mathbf{\Phi}(t_2)\mathbf{\Phi}(t_1)$
4. $[\mathbf{\Phi}(t)]^n = \mathbf{\Phi}(nt)$
5. $\mathbf{\Phi}(t_2 - t_1)\mathbf{\Phi}(t_1 - t_0) = \mathbf{\Phi}(t_2 - t_0) = \mathbf{\Phi}(t_1 - t_0)\mathbf{\Phi}(t_2 - t_1)$

Exemplo 14.6 Obtenha a matriz de transição de estados $\mathbf{\Phi}(t)$ do seguinte sistema:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Obtenha também a inversa da matriz de transição de estados, $\mathbf{\Phi}^{-1}(t)$.

Para este sistema,

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$$

A matriz de transição de estado $\mathbf{\Phi}(t)$ é dada por

$$\mathbf{\Phi}(t) = e^{\mathbf{A}t} = \mathscr{L}^{-1}[(s\mathbf{I} - \mathbf{A})^{-1}]$$

como

$$s\mathbf{I} - \mathbf{A} = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} = \begin{bmatrix} s & -1 \\ 2 & s+3 \end{bmatrix}$$

A inversa de $(s\mathbf{I} - \mathbf{A})$ é dada por

$$(s\mathbf{I} - \mathbf{A})^{-1} = \frac{1}{(s+1)(s+2)} \begin{bmatrix} s+3 & 1 \\ -2 & s \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{s+3}{(s+1)(s+2)} & \dfrac{1}{(s+1)(s+2)} \\ \dfrac{-2}{(s+1)(s+2)} & \dfrac{s}{(s+1)(s+2)} \end{bmatrix}$$

Portanto

$$\mathbf{\Phi}(t) = e^{\mathbf{A}t} = \mathscr{L}^{-1}[(s\mathbf{I} - \mathbf{A})^{-1}]$$

$$= \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix}$$

Observando que $\mathbf{\Phi}^{-1}(t) = \mathbf{\Phi}(-t)$, obtemos a inversa da matriz de transição de estados como segue:

$$\mathbf{\Phi}^{-1}(t) = e^{-\mathbf{A}t} = \begin{bmatrix} 2e^{t} - e^{2t} & e^{t} - e^{2t} \\ -2e^{t} + 2e^{2t} & -e^{t} + 2e^{2t} \end{bmatrix}$$

Solução das equações de estado não homogêneas. Iniciaremos considerando o caso escalar

$$\dot{x} = ax + bu \tag{14.43}$$

Vamos reescrever a Eq. (14.43) como

$$\dot{x} - ax = bu$$

Multiplicando ambos os lados desta equação por e^{-at}, obtemos

$$e^{-at}[\dot{x}(t) - ax(t)] = \frac{d}{dt}[e^{-at}x(t)] = e^{-at}bu(t)$$

Integrando esta equação entre 0 e t, resulta

$$e^{-at}x(t) = x(0) + \int_0^t e^{-a\tau}bu(\tau)\,d\tau$$

ou

$$x(t) = e^{at}x(0) + e^{at}\int_0^t e^{-a\tau}bu(\tau)\,d\tau$$

O primeiro termo do lado direito é a resposta a condições iniciais, e o segundo termo é a resposta à entrada $u(t)$.

Vamos agora considerar a equação de estados não homogênea descrita por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{14.44}$$

onde

$\mathbf{x}$ = vetor n-dimensional
$\mathbf{u}$ = vetor r-dimensional
$\mathbf{A}$ = matriz constante $n \times n$
$\mathbf{B}$ = matriz constante $n \times r$

Escrevendo a Eq. (14.44) como

$$\dot{\mathbf{x}}(t) - \mathbf{A}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t)$$

e pré-multiplicando ambos os lados desta equação por $e^{-\mathbf{A}t}$, obtemos

$$e^{-\mathbf{A}t}[\dot{\mathbf{x}}(t) - \mathbf{A}\mathbf{x}(t)] = \frac{d}{dt}[e^{-\mathbf{A}t}\mathbf{x}(t)] = e^{-\mathbf{A}t}\mathbf{B}\mathbf{u}(t)$$

Integrando a equação anterior entre 0 e t, resulta

$$e^{-\mathbf{A}t}\mathbf{x}(t) = \mathbf{x}(0) + \int_0^t e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)\,d\tau$$

ou

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau \tag{14.45}$$

A Eq. (14.45) pode também ser escrita como

$$\mathbf{x}(t) = \mathbf{\Phi}(t)\mathbf{x}(0) + \int_0^t \mathbf{\Phi}(t-\tau)\mathbf{B}\mathbf{u}(\tau)\,d\tau \tag{14.46}$$

onde

$$\mathbf{\Phi}(t) = e^{\mathbf{A}t}$$

A Eq. (14.45) ou (14.46) é a solução da Eq. (14.44). A solução $\mathbf{x}(t)$ é claramente a soma de um termo que consiste na transição do estado inicial e um termo proveniente do vetor de entrada.

Abordagem por transformada de Laplace para a solução de equações de estado não homogêneas. A solução da equação de estado não homogênea

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

pode também ser obtida pela abordagem por transformada de Laplace. A transformada de Laplace da Eq. (14.44) resulta

$$s\mathbf{X}(s) - \mathbf{x}(0) = \mathbf{A}\mathbf{X}(s) + \mathbf{B}\mathbf{U}(s)$$

ou

$$(s\mathbf{I} - \mathbf{A})\mathbf{X}(s) = \mathbf{x}(0) + \mathbf{B}\mathbf{U}(s)$$

Pré-multiplicando ambos os lados desta última equação por $(s\mathbf{I} - \mathbf{A})^{-1}$, obtemos

$$\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0) + (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{U}(s)$$

Usando esta relação dada pela Eq. (14.40), resulta

$$\mathbf{X}(s) = \mathscr{L}[e^{\mathbf{A}t}]\mathbf{x}(0) + \mathscr{L}[e^{\mathbf{A}t}]\mathbf{B}\mathbf{U}(s)$$

A transformada inversa de Laplace desta última equação pode ser obtida usando-se a integral de convolução como segue:

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau$$

Solução em termos de $x(t_0)$. Até agora supusemos que o tempo inicial é zero. Entretanto, se o tempo inicial é dado por t_0 ao invés de 0, então a solução da Eq. (14.44) deve ser modificada para

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau \tag{14.47}$$

Exemplo 14.7 Obtenha a resposta temporal do seguinte sistema:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}0 & 1\\ -2 & -3\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix}[u]$$

onde $u(t)$ é a função degrau unitário ocorrendo em $t = 0$, ou

$$u(t) = 1(t)$$

Para este sistema,

$$\mathbf{A} = \begin{bmatrix}0 & 1\\ -2 & -3\end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix}0\\ 1\end{bmatrix}$$

A matriz de transição de estados $\mathbf{\Phi}(t) = e^{\mathbf{A}t}$ foi obtida no Exemplo 14.6 como

$$\mathbf{\Phi}(t) = e^{\mathbf{A}t} = \begin{bmatrix}2e^{-t} - e^{-2t} & e^{-t} - e^{-2t}\\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t}\end{bmatrix}$$

A resposta à entrada degrau unitário é então obtida como

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t \begin{bmatrix} 2e^{-(t-\tau)} - e^{-2(t-\tau)} & e^{-(t-\tau)} - e^{-2(t-\tau)} \\ -2e^{-(t-\tau)} + 2e^{-2(t-\tau)} & -e^{-(t-\tau)} + 2e^{-2(t-\tau)} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} [1]\, d\tau$$

ou

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + \begin{bmatrix} \frac{1}{2} - e^{-t} + \frac{1}{2}e^{-2t} \\ e^{-t} - e^{-2t} \end{bmatrix}$$

Se o estado inicial é zero, ou $\mathbf{x}(0) = \mathbf{0}$, então $\mathbf{x}(t)$ pode ser simplificado para

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} - e^{-t} + \frac{1}{2}e^{-2t} \\ e^{-t} - e^{-2t} \end{bmatrix}$$

14.4 MATRIZ DE TRANSFERÊNCIA

Na Seção 4.6 definimos a matriz de transferência. Como o conceito da matriz de transferência é uma extensão da função de transferência, inicialmente obteremos funções de transferência de sistemas de uma entrada e uma saída e então matrizes de transferência de sistemas de múltiplas entradas e múltiplas saídas a partir de equações de estado e de saída.

Funções de transferência. Derivaremos a função de transferência de um sistema de uma entrada e uma saída a partir da versão transformada segundo Laplace das equações de estado e de saída.

Vamos considerar o sistema cuja função de transferência é dada por

$$\frac{Y(s)}{U(s)} = G(s) \tag{14.48}$$

Na Seção 14.3, mostramos que a representação de espaço de estados deste sistema é dada por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \tag{14.49}$$

$$y = \mathbf{C}\mathbf{x} + Du \tag{14.50}$$

onde $\mathbf{x}$ é o vetor de estado, u é a entrada, e y é a saída. As transformadas de Laplace das Eqs. (14.49) e (14.50) são dadas por

$$s\mathbf{X}(s) - \mathbf{x}(0) = \mathbf{A}\mathbf{X}(s) + \mathbf{B}U(s) \tag{14.51}$$

$$Y(s) = \mathbf{C}\mathbf{X}(s) + DU(s) \tag{14.52}$$

Como a função de transferência foi anteriormente definida como a razão da transformada de Laplace da saída para a transformada de Laplace da entrada quando as condições iniciais são zero, supomos que $\mathbf{x}(0)$ na Eq. (14.51) é zero.

Substituindo $\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}U(s)$ na Eq. (14.52), obtemos

$$Y(s) = [\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + D]U(s) \tag{14.53}$$

Comparando a Eq. (14.53) com a Eq. (14.48), vemos que

$$G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + D \tag{14.54}$$

Esta é a expressão da função de transferência em termos de **A**, **B**, **C** e D.

Observe que o lado direito da Eq. (14.54) envolve $(s\mathbf{I} - \mathbf{A})^{-1}$. Portanto $G(s)$ pode ser escrito como

$$G(s) = \frac{Q(s)}{|s\mathbf{I} - \mathbf{A}|}$$

onde $Q(s)$ é um polinômio em s. Portanto, $|s\mathbf{I} - \mathbf{A}|$ é igual ao polinômio característico de $G(s)$. Em outras palavras, os autovalores de **A** são idênticos aos pólos de $G(s)$.

Exemplo 14.8 Obtenha a função de transferência do sistema visto na Fig. 14.7.

Do diagrama, obtemos as seguintes equações de estado e de saída:

$$\dot{x}_1 = -5x_1 - x_2 + 2u$$
$$\dot{x}_2 = 3x_1 - x_2 + 5u$$
$$y = x_1 + 2x_2$$

Na forma matricial, temos

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -5 & -1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} [u]$$

$$y = [1 \quad 2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

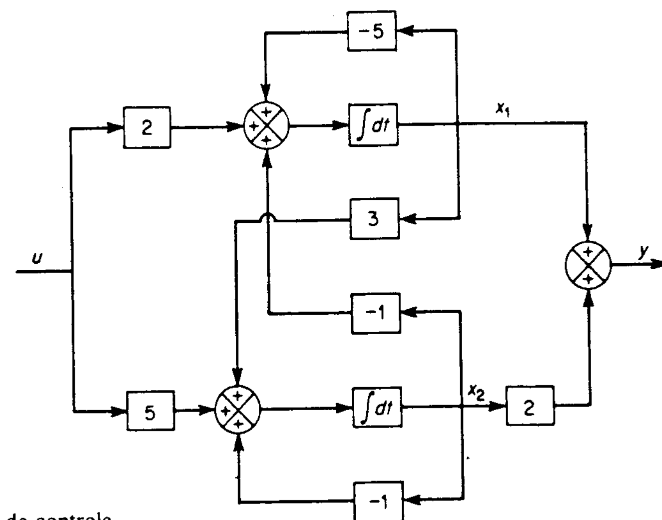


Fig. 14.7 Sistema de controle.

A função de transferência para o sistema é então

$$
\begin{aligned}
G(s) &= \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \\
&= [1 \quad 2]\begin{bmatrix} s+5 & 1 \\ -3 & s+1 \end{bmatrix}^{-1}\begin{bmatrix} 2 \\ 5 \end{bmatrix} \\
&= [1 \quad 2]\begin{bmatrix} \dfrac{s+1}{(s+2)(s+4)} & \dfrac{-1}{(s+2)(s+4)} \\ \dfrac{3}{(s+2)(s+4)} & \dfrac{s+5}{(s+2)(s+4)} \end{bmatrix}\begin{bmatrix} 2 \\ 5 \end{bmatrix} \\
&= \frac{12s+59}{(s+2)(s+4)}
\end{aligned}
$$

Matriz de transferência. A matriz de transferência $\mathbf{G}(s)$ relaciona a saída $\mathbf{Y}(s)$ com a entrada $\mathbf{U}(s)$, ou

$$\mathbf{Y}(s) = \mathbf{G}(s)\mathbf{U}(s) \tag{14.55}$$

Se o vetor de entrada $\mathbf{u}$ é r-dimensional e o vetor de saída $\mathbf{y}$ é m-dimensional, então a matriz de transferência é uma matriz $m \times r$. Em uma forma expandida, a Eq. (14.55) pode ser escrita como

$$
\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix} = \begin{bmatrix} G_{11}(s) & G_{12}(s) & \cdots & G_{1r}(s) \\ G_{21}(s) & G_{22}(s) & \cdots & G_{2r}(s) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ G_{m1}(s) & G_{m2}(s) & \cdots & G_{mr}(s) \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_r \end{bmatrix}
$$

O elemento de ordem (i, j) $G_{ij}(s)$ de $\mathbf{G}(s)$ é a função de transferência relacionando a i-ésima saída com a j-ésima entrada.

Seguindo os mesmos passos usados na derivação da Eq. (14.54), obtemos a matriz de transferência para sistemas de múltiplas entradas e múltiplas saídas como segue:

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$$

Claramente, a expressão da função de transferência dada pela Eq. (14.54) é um caso especial desta expressão da matriz de transferência.

Matriz de transferência de sistemas a malha-fechada. Considere o sistema visto na Fig. 14.8. O sistema tem múltiplas entradas e múltiplas saídas. A matriz de transferência do elo direto é $\mathbf{G}_0(s)$, e a do elo de realimentação é $\mathbf{H}(s)$. A matriz de transferência entre o vetor de sinal de realimentação $\mathbf{B}(s)$ e o vetor de erro $\mathbf{E}(s)$ é obtida como segue: Como

$$
\begin{aligned}
\mathbf{B}(s) &= \mathbf{H}(s)\mathbf{Y}(s) \\
&= \mathbf{H}(s)\mathbf{G}_0(s)\mathbf{E}(s)
\end{aligned}
$$

obtemos a matriz de transferência entre $\mathbf{B}(s)$ e $\mathbf{E}(s)$ como $\mathbf{H}(s)\mathbf{G}_0(s)$. Portanto a

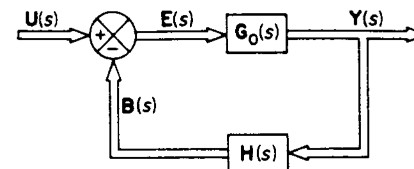


Fig. 14.8 Diagrama de blocos de um sistema de múltiplas entradas e múltiplas saídas.

matriz de transferência dos elementos em cascata é o produto das matrizes de transferência dos elementos individuais. (Note que a ordem da multiplicação matricial é muito importante, uma vez que em geral a multiplicação matricial não é comutativa.)

A matriz de transferência do sistema a malha-fechada é obtida como segue: Como

$$
\begin{aligned}
\mathbf{Y}(s) &= \mathbf{G}_0(s)[\mathbf{U}(s) - \mathbf{B}(s)] \\
&= \mathbf{G}_0(s)[\mathbf{U}(s) - \mathbf{H}(s)\mathbf{Y}(s)]
\end{aligned}
$$

obtemos

$$[\mathbf{I} + \mathbf{G}_0(s)\mathbf{H}(s)]\mathbf{Y}(s) = \mathbf{G}_0(s)\mathbf{U}(s)$$

Pré-multiplicando ambos os lados desta última equação por $[\mathbf{I} + \mathbf{G}_0(s)\mathbf{H}(s)]^{-1}$, obtemos

$$\mathbf{Y}(s) = [\mathbf{I} + \mathbf{G}_0(s)\mathbf{H}(s)]^{-1}\mathbf{G}_0(s)\mathbf{U}(s)$$

A matriz de transferência de malha-fechada $\mathbf{G}(s)$ é então dada por

$$\mathbf{G}(s) = [\mathbf{I} + \mathbf{G}_0(s)\mathbf{H}(s)]^{-1}\mathbf{G}_0(s) \tag{14.56}$$

Não-interação em sistemas de múltiplas entradas e saídas. Muitos sistemas de controle de processos têm múltiplas entradas e múltiplas saídas, e freqüentemente se deseja que mudanças em uma entrada de referência afetem apenas uma saída. (Se podemos obter tal não-interação, é mais fácil manter cada valor de saída em um valor constante desejado na ausência de perturbações externas.)

Vamos considerar a matriz de transferência $\mathbf{G}_p(s)$ (uma matriz $n \times n$) de um processo e projetar um compensador série $\mathbf{G}_c(s)$ (também uma matriz $n \times n$) tal que as n entradas e n saídas são desacopladas. Se desejarmos a não-interação ou desacoplamento entre as n entradas e n saídas, a matriz de transferência de malha-fechada deve ser diagonal, ou

$$
\mathbf{G}(s) = \begin{bmatrix} G_{11}(s) & & & & 0 \\ & G_{22}(s) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ 0 & & & & G_{nn}(s) \end{bmatrix}
$$

Consideraremos o caso em que a matriz de realimentação $\mathbf{H}(s)$ é a matriz identidade. Então, da Eq. (14.56), obtemos

$$\mathbf{G}(s) = [\mathbf{I} + \mathbf{G}_0(s)]^{-1}\mathbf{G}_0(s) \tag{14.57}$$

onde

$$\mathbf{G}_0(s) = \mathbf{G}_p(s)\mathbf{G}_c(s)$$

Da Eq. (14.57), obtemos

$$[\mathbf{I} + \mathbf{G}_0(s)]\mathbf{G}(s) = \mathbf{G}_0(s)$$

ou

$$\mathbf{G}_0(s)[\mathbf{I} - \mathbf{G}(s)] = \mathbf{G}(s)$$

Pós-multiplicando ambos os lados desta última equação por $[\mathbf{I} - \mathbf{G}(s)]^{-1}$, obtemos

$$\mathbf{G}_0(s) = \mathbf{G}(s)[\mathbf{I} - \mathbf{G}(s)]^{-1}$$

Como $\mathbf{G}(s)$ é uma matriz diagonal, $\mathbf{I} - \mathbf{G}(s)$ é também uma matriz diagonal. Então $\mathbf{G}_0(s)$, um produto de duas matrizes diagonais, também é uma matriz diagonal. Isto significa que, para se conseguir não-interação, devemos fazer $\mathbf{G}_0(s)$ uma matriz diagonal, contanto que a matriz de realimentação $\mathbf{H}(s)$ seja a matriz identidade.

Exemplo 14.9 Considere o sistema visto na Fig. 14.9. Determine a matriz de transferência do compensador série para que a matriz de transferência de malha-fechada seja

$$\mathbf{G}(s) = \begin{bmatrix} \frac{1}{s+1} & 0 \\ 0 & \frac{1}{5s+1} \end{bmatrix}$$

como

$$\mathbf{G}_0 = \mathbf{G}(\mathbf{I} - \mathbf{G})^{-1}$$

$$= \begin{bmatrix} \frac{1}{s+1} & 0 \\ 0 & \frac{1}{5s+1} \end{bmatrix} \begin{bmatrix} \frac{s+1}{s} & 0 \\ 0 & \frac{5s+1}{5s} \end{bmatrix} = \begin{bmatrix} \frac{1}{s} & 0 \\ 0 & \frac{1}{5s} \end{bmatrix}$$

$$\begin{bmatrix} Y_1(s) \\ Y_2(s) \end{bmatrix} = \begin{bmatrix} \frac{1}{2s+1} & 0 \\ 1 & \frac{1}{s-1} \end{bmatrix} \begin{bmatrix} U_1(s) \\ U_2(s) \end{bmatrix}$$

e

$$\begin{bmatrix} U_1(s) \\ U_2(s) \end{bmatrix} = \begin{bmatrix} G_{c11}(s) & G_{c12}(s) \\ G_{c21}(s) & G_{c22}(s) \end{bmatrix} \begin{bmatrix} R_1(s) - Y_1(s) \\ R_2(s) - Y_2(s) \end{bmatrix}$$

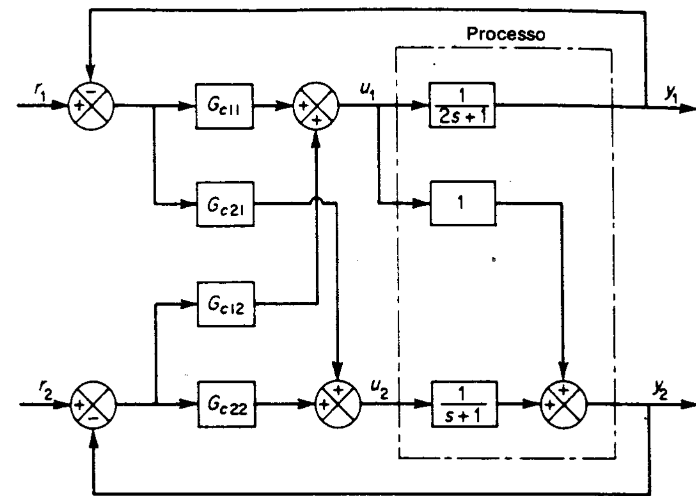


Fig. 14.9 Sistema de múltiplas entradas e múltiplas saídas.

obtemos

$$\begin{bmatrix} Y_1(s) \\ Y_2(s) \end{bmatrix} = \begin{bmatrix} \frac{1}{2s+1} & 0 \\ 1 & \frac{1}{s+1} \end{bmatrix} \begin{bmatrix} G_{c11}(s) & G_{c12}(s) \\ G_{c21}(s) & G_{c22}(s) \end{bmatrix} \begin{bmatrix} R_1(s) - Y_1(s) \\ R_2(s) - Y_2(s) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{s} & 0 \\ 0 & \frac{1}{5s} \end{bmatrix} \begin{bmatrix} R_1(s) - Y_1(s) \\ R_2(s) - Y_2(s) \end{bmatrix}$$

Portanto

$$\mathbf{G}_c(s) = \begin{bmatrix} G_{c11}(s) & G_{c12}(s) \\ G_{c21}(s) & G_{c22}(s) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2s+1} & 0 \\ 1 & \frac{1}{s+1} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{s} & 0 \\ 0 & \frac{1}{5s} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{2s+1}{s} & 0 \\ -\frac{(s+1)(2s+1)}{s} & \frac{s+1}{5s} \end{bmatrix} \tag{14.58}$$

A Eq. (14.58) nos dá a matriz de transferência do compensador série. Note que $G_{c11}(s)$ e $G_{c22}(s)$ são controladores do tipo proporcional-mais-integral e $G_{c21}(s)$ é um controlador do tipo proporcional-mais-integral-mais derivada.

É muito importante notar que na análise presente não consideramos perturbações externas. Em geral, na presente abordagem, há cancelamentos no numerador e denominador. Portanto, alguns dos autovalores serão perdidos em $G_p(s)G_c(s)$. Isto significa que embora a abordagem presente forneça o resultado desejado de não-interação nas respostas a entradas de referência na ausência de perturbações externas. caso o sistema seja perturbado por forças externas, então o sistema pode-se tornar incontrolável porque qualquer movimento causado pelo autovalor cancelado não pode ser controlado. (Discutiremos detalhes de controlabilidade de sistemas na Seção 16.2.)

14.5 SISTEMAS LINEARES VARIANTES NO TEMPO

Uma vantagem do método de espaço de estados na análise de sistemas de controle é que ele pode ser estendido a sistemas lineares variantes no tempo.

A maioria dos resultados obtidos na Seção 14.4 valem para sistemas lineares variantes no tempo modificando-se a matriz de transição $\mathbf{\Phi}(t)$ para $\mathbf{\Phi}(t, t_0)$. (Para sistemas variantes no tempo, a matriz de transição depende tanto de t como de t_0 e não da diferença $t - t_0$. Portanto, não podemos sempre fazer o tempo inicial igual a zero. Há, entretanto, casos em que t_0 é zero.) Entretanto, é importante compreender que a matriz de transição para um sistema variante no tempo não pode em geral ser dada como uma matriz exponencial.

Solução de equações de estado variantes no tempo. Para uma equação diferencial escalar

$$\dot{x} = a(t)x$$

a solução pode ser dada por

$$x(t) = e^{\int_{t_0}^{t} a(\tau)\, d\tau} x(t_0)$$

e a função de transição de estados pode ser dada por

$$\phi(t, t_0) = \exp\left[\int_{t_0}^{t} a(\tau)\, d\tau\right]$$

O mesmo resultado, entretanto, não se aplica para a equação diferencial matricial.

Considere a equação de estado

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} \tag{14.59}$$

onde

$\mathbf{x}(t)$ = vetor n-dimensional

$\mathbf{A}(t)$ = matriz $n \times n$ cujos elementos são funções contínuas por trechos em t no intervalo $t_0 \leq t \leq t_1$

A solução da Eq. (14.59) é dada por

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0)\mathbf{x}(t_0) \tag{14.60}$$

onde $\mathbf{\Phi}(t, t_0)$ é a matriz não singular $n \times n$ que satisfaz a seguinte equação diferencial

$$\dot{\mathbf{\Phi}}(t, t_0) = \mathbf{A}(t)\mathbf{\Phi}(t, t_0), \qquad \mathbf{\Phi}(t_0, t_0) = \mathbf{I} \tag{14.61}$$

O fato de que a Eq. (14.60) é a solução da Eq. (14.61) pode ser verificado facilmente pois

$$\mathbf{x}(t_0) = \mathbf{\Phi}(t_0, t_0)\mathbf{x}(t_0) = \mathbf{I}\mathbf{x}(t_0)$$

e

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \frac{d}{dt}[\mathbf{\Phi}(t, t_0)\mathbf{x}(t_0)] \\ &= \dot{\mathbf{\Phi}}(t, t_0)\mathbf{x}(t_0) \\ &= \mathbf{A}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}(t_0) = \mathbf{A}(t)\mathbf{x}(t)\end{aligned}$$

Vemos que a solução da Eq. (14.59) é simplesmente uma transformação do estado inicial. A matriz $\mathbf{\Phi}(t, t_0)$ é a matriz de transição de estados para o sistema variante no tempo descrito pela Eq. (14.59).

Matriz de transição de estado para o caso variante no tempo. É importante notar que a matriz de transição de estados $\mathbf{\Phi}(t, t_0)$ é dada por uma matriz exponencial se e apenas se $\mathbf{A}(t)$ e $\int_{t_0}^{t} \mathbf{A}(\tau)d\tau$ comutam. Isto é,

$$\mathbf{\Phi}(t, t_0) = \exp\left[\int_{t_0}^{t} \mathbf{A}(\tau)\, d\tau\right] \quad \text{(Se e apenas se } \mathbf{A}(t) \text{ e } \int_{t_0}^{t} \mathbf{A}(\tau)\, d\tau \text{ comutam.)}$$

Note que se $\mathbf{A}(t)$ é uma matriz constante ou uma matriz diagonal, $\mathbf{A}(t)$ e $\int_{t_0}^{t} \mathbf{A}(\tau)d\tau$ comutam. Se $\mathbf{A}(t)$ e $\int_{t_0}^{t} \mathbf{A}(\tau)d\tau$ não comutam, não há uma maneira simples de se computar a matriz de transição de estados.

Para computar $\mathbf{\Phi}(t, t_0)$ numericamente, podemos usar a seguinte expansão em série para $\mathbf{\Phi}(t, t_0)$:

$$\mathbf{\Phi}(t, t_0) = \mathbf{I} + \int_{t_0}^{t} \mathbf{A}(\tau)\, d\tau + \int_{t_0}^{t} \mathbf{A}(\tau_1)\left[\int_{t_0}^{\tau_1} \mathbf{A}(\tau_2)\, d\tau_2\right] d\tau_1 + \cdots \tag{14.62}$$

Isto, em geral, não fornecerá $\mathbf{\Phi}(t, t_0)$ em forma fechada.

Exemplo 14.10 Obtenha $\mathbf{\Phi}(t, 0)$ para o sistema variante no tempo

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & t\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

Para computar $\mathbf{\Phi}(t, 0)$, vamos usar a Eq. (14.62). Como

$$\int_0^t \mathbf{A}(\tau)\, d\tau = \int_0^t \begin{bmatrix}0 & 1\\ 0 & \tau\end{bmatrix} d\tau = \begin{bmatrix}0 & t\\ 0 & \frac{t^2}{2}\end{bmatrix}$$

$$\int_0^t \begin{bmatrix}0 & 1\\ 0 & \tau_1\end{bmatrix}\left\{\int_0^{\tau_1} \begin{bmatrix}0 & 1\\ 0 & \tau_2\end{bmatrix} d\tau_2\right\} d\tau_1 = \int_0^t \begin{bmatrix}0 & 1\\ 0 & \tau_1\end{bmatrix}\begin{bmatrix}0 & \tau_1\\ 0 & \frac{\tau_1^2}{2}\end{bmatrix} d\tau_1 = \begin{bmatrix}0 & \frac{t^3}{6}\\ 0 & \frac{t^4}{8}\end{bmatrix}$$

obtemos

$$\mathbf{\Phi}(t,0)=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\begin{bmatrix}0 & t\\ 0 & \frac{t^2}{2}\end{bmatrix}+\begin{bmatrix}0 & \frac{t^3}{6}\\ 0 & \frac{t^4}{8}\end{bmatrix}+\cdots$$

$$=\begin{bmatrix}1 & t+\frac{t^3}{6}+\cdots\\ 0 & 1+\frac{t^2}{2}+\frac{t^4}{8}+\cdots\end{bmatrix}$$

Propriedades da matriz de transição de estados $\Phi(t, t_0)$. A seguir alistaremos as propriedades da matriz de transição de estados $\Phi(t, t_0)$.

1. $\mathbf{\Phi}(t_2, t_1)\mathbf{\Phi}(t_1, t_0) = \mathbf{\Phi}(t_2, t_0)$

Para provar esta igualdade, note que

$$\mathbf{x}(t_1) = \mathbf{\Phi}(t_1, t_0)\mathbf{x}(t_0)$$
$$\mathbf{x}(t_2) = \mathbf{\Phi}(t_2, t_0)\mathbf{x}(t_0)$$

Também

$$\mathbf{x}(t_2) = \mathbf{\Phi}(t_2, t_1)\mathbf{x}(t_1)$$

Portanto

$$\mathbf{x}(t_2) = \mathbf{\Phi}(t_2, t_1)\mathbf{\Phi}(t_1, t_0)\mathbf{x}(t_0) = \mathbf{\Phi}(t_2, t_0)\mathbf{x}(t_0)$$

Portanto

$$\mathbf{\Phi}(t_2, t_1)\mathbf{\Phi}(t_1, t_0) = \mathbf{\Phi}(t_2, t_0)$$

2. $\mathbf{\Phi}(t_1, t_0) = \mathbf{\Phi}^{-1}(t_0, t_1)$

Para provar isto, note que

$$\mathbf{\Phi}(t_1, t_0) = \mathbf{\Phi}^{-1}(t_2, t_1)\mathbf{\Phi}(t_2, t_0)$$

Se fizermos $t_2 = t_0$ nesta última equação, então

$$\mathbf{\Phi}(t_1, t_0) = \mathbf{\Phi}^{-1}(t_0, t_1)\mathbf{\Phi}(t_0, t_0) = \mathbf{\Phi}^{-1}(t_0, t_1)$$

Solução de equações de estado lineares variantes no tempo. Considere a seguinte equação de estado:

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u} \tag{14.63}$$

onde

$\mathbf{x}$ = vetor n-dimensional
$\mathbf{u}$ = vetor r-dimensional

$\mathbf{A}(t)$ = matriz $n \times n$
$\mathbf{B}(t)$ = matriz $n \times r$

Supõe-se que os elementos de $\mathbf{A}(t)$ e $\mathbf{B}(t)$ são funções contínuas por trechos de t no intervalo $t_0 \leq t \leq t_1$.

Para se obter a solução da Eq. (14.63), vamos fazer

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0)\boldsymbol{\xi}(t)$$

onde $\mathbf{\Phi}(t, t_0)$ é a matriz única satisfazendo a seguinte equação:

$$\dot{\mathbf{\Phi}}(t, t_0) = \mathbf{A}(t)\mathbf{\Phi}(t, t_0),$$
$$\mathbf{\Phi}(t_0, t_0) = \mathbf{I}$$

Então

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \frac{d}{dt}[\mathbf{\Phi}(t, t_0)\boldsymbol{\xi}(t)]\\ &= \dot{\mathbf{\Phi}}(t, t_0)\boldsymbol{\xi}(t) + \mathbf{\Phi}(t, t_0)\dot{\boldsymbol{\xi}}(t)\\ &= \mathbf{A}(t)\mathbf{\Phi}(t, t_0)\boldsymbol{\xi}(t) + \mathbf{\Phi}(t, t_0)\dot{\boldsymbol{\xi}}(t)\\ &= \mathbf{A}(t)\mathbf{\Phi}(t, t_0)\boldsymbol{\xi}(t) + \mathbf{B}(t)\mathbf{u}(t)\end{aligned}$$

Portanto

$$\mathbf{\Phi}(t, t_0)\dot{\boldsymbol{\xi}}(t) = \mathbf{B}(t)\mathbf{u}(t)$$

ou

$$\dot{\boldsymbol{\xi}}(t) = \mathbf{\Phi}^{-1}(t, t_0)\mathbf{B}(t)\mathbf{u}(t)$$

Portanto

$$\boldsymbol{\xi}(t) = \boldsymbol{\xi}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}^{-1}(\tau, t_0)\mathbf{B}(\tau)\mathbf{u}(\tau)\,d\tau$$

Como

$$\boldsymbol{\xi}(t_0) = \mathbf{\Phi}^{-1}(t_0, t_0)\mathbf{x}(t_0) = \mathbf{x}(t_0)$$

a solução da Eq. (14.63) é obtida como

$$\begin{aligned}\mathbf{x}(t) &= \mathbf{\Phi}(t, t_0)\mathbf{x}(t_0) + \mathbf{\Phi}(t, t_0)\int_{t_0}^{t} \mathbf{\Phi}^{-1}(\tau, t_0)\mathbf{B}(\tau)\mathbf{u}(\tau)\,d\tau\\ &= \mathbf{\Phi}(t, t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\,d\tau\end{aligned} \tag{14.64}$$

O cálculo do lado direito da Eq. (14.64) para casos práticos requer um computador digital.

14.6 REPRESENTAÇÃO DE SISTEMAS DE TEMPO DISCRETO POR ESPAÇO DE ESTADOS

A abordagem de espaço de estados para a análise de sistemas dinâmicos pode ser estendida para o caso de tempo discreto. A forma discreta da representação de espaço de estados é bastante análoga à forma contínua.

A representação de espaço de estados mais geral para sistemas lineares de tempo discreto é

$$\mathbf{x}(k+1) = \mathbf{G}(k)\mathbf{x}(k) + \mathbf{H}(k)\mathbf{u}(k) \tag{14.65}$$

$$\mathbf{y}(k) = \mathbf{C}(k)\mathbf{x}(k) + \mathbf{D}(k)\mathbf{u}(k) \tag{14.66}$$

onde $\mathbf{x}(k)$ é o vetor de estados, $\mathbf{u}(k)$ é o vetor de entrada, e $\mathbf{y}(k)$ é o vetor de saída, cada um especificado em $t = kT$, $k = 0, 1, 2, \ldots$, e T é o período de amostragem. [Note que a menos que especifiquemos em contrário, usaremos a notação simplificada $\mathbf{x}(k)$ para indicar $\mathbf{x}(kT)$. Isto é, $\mathbf{x}(k)$ implica o vetor $\mathbf{x}(t)$ em $t = kT$. De forma similar, usamos a notação simplificada $\mathbf{u}(k)$, $\mathbf{y}(k)$, $\mathbf{G}(k)$, $\mathbf{H}(k)$, $\mathbf{C}(k)$ e $\mathbf{D}(k)$.] As Eqs. (14.65) e (14.66) correspondem ao caso variante no tempo. A Fig. 14.10 mostra o diagrama de blocos do sistema de tempo discreto descrito pelas Eqs. (14.65) e (14.66). O elemento atraso unitário tem um tempo de atraso de T segundos.

Se o sistema linear de tempo discreto é invariante no tempo, então as Eqs. (14.65) e (14.66) são modificadas para

$$\mathbf{x}(k+1) = \mathbf{G}\mathbf{x}(k) + \mathbf{H}\mathbf{u}(k) \tag{14.67}$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \tag{14.68}$$

Nesta seção, estaremos interessados principalmente com sistemas descritos pelas Eqs. (14.67) e (14.68).

Representação por espaço de estados de equações de diferenças escalares e invariantes no tempo onde a função de excitação é $bu(k)$. Considere a seguinte

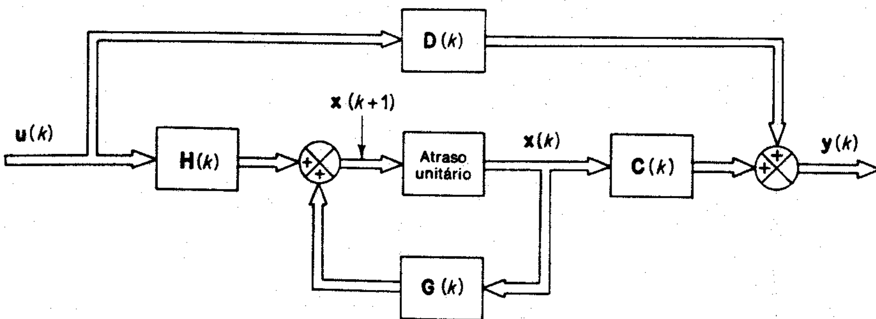


Fig. 14.10 Representação em diagrama de blocos do sistema de tempo discreto descrito pelas Eqs. (14.65) e (14.66).

equação de diferenças escalar:

$$\begin{aligned} y(k+n) + a_1 y(k+n-1) + a_2 y(k+n-2) + \cdots \\ + a_{n-1} y(k+1) + a_n y(k) = bu(k) \end{aligned} \tag{14.69}$$

onde k denota o k-ésimo instante de amostragem, $y(k)$ é a saída do sistema no k-ésimo instante de amostragem, e $u(k)$ é a entrada no k-ésimo instante de amostragem. Vamos definir

$$\begin{aligned} x_1(k) &= y(k) \\ x_1(k+1) &= x_2(k) \\ x_2(k+1) &= x_3(k) \\ &\cdots \\ x_{n-1}(k+1) &= x_n(k) \\ x_n(k+1) &= -a_1 x_n(k) - a_2 x_{n-1}(k) - \cdots - a_n x_1(k) + bu(k) \end{aligned}$$

Então, a Eq. (14.69) pode ser escrita na seguinte forma:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1}(k+1) \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1}(k) \\ x_n(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ b \end{bmatrix} [u(k)]$$

$$y(k) = [1 \quad 0 \quad \cdots \quad 0] \begin{bmatrix} x_1(k) \\ x_2(k) \\ \cdot \\ \cdot \\ \cdot \\ x_n(k) \end{bmatrix}$$

ou

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{G}\mathbf{x}(k) + \mathbf{H}u(k) \\ y(k) &= \mathbf{C}\mathbf{x}(k) \end{aligned}$$

onde

$$\mathbf{x}(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_{n-1}(k) \\ x_n(k) \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b \end{bmatrix}$$

$$\mathbf{C} = [1 \quad 0 \quad \cdots \quad 0 \quad 0]$$

Representação de espaço de estados de equações de diferenças escalares invariantes no tempo quando a função de excitado envolve $u(k), u(k+1), \ldots, u(k+n)$. Considere em seguida a seguinte equação de diferenças escalar:

$$y(k+n) + a_1 y(k+n-1) + a_2 y(k+n-2) + \cdots + a_{n-1} y(k+1)$$
$$+ a_n y(k) = b_0 u(k+n) + b_1 u(k+n-1) + b_2 u(k+n-2)$$
$$+ \cdots + b_{n-1} u(k+1) + b_n u(k) \tag{14.70}$$

onde k denota o k-ésimo instante de amostragem, $y(k)$ é a saída do sistema no k-ésimo instante de amostragem, e $u(k)$ é a entrada no k-ésimo instante de amostragem.

De forma semelhante ao caso do sistema de equações diferenciais escalares dadas pela Eq. (14.24), definamos as variáveis de estado como segue:

$$x_1(k) = y(k) - h_0 u(k)$$
$$x_2(k) = x_1(k+1) - h_1 u(k)$$
$$x_3(k) = x_2(k+1) - h_2 u(k)$$
$$\cdots$$
$$x_n(k) = x_{n-1}(k+1) - h_{n-1} u(k)$$

onde $h_0, h_1, h_2, \ldots, h_n$ são determinadas de

$$h_0 = b_0$$
$$h_1 = b_1 - a_1 h_0$$
$$h_2 = b_2 - a_1 h_1 - a_2 h_0$$
$$\cdots$$
$$h_n = b_n - a_1 h_{n-1} - \cdots - a_{n-1} h_1 - a_n h_0$$

Com esta escolha das variáveis de estado, obtemos a seguinte equação de estado de tempo discreto e equação de saída para o sistema da Eq. (14.70):

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_{n-1}(k+1) \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_{n-1}(k) \\ x_n(k) \end{bmatrix} +$$

$$+ \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_{n-1} \\ h_n \end{bmatrix} [u(k)]$$

$$y(k) = [1 \quad 0 \quad \cdots \quad 0] \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{bmatrix} + h_0 u(k)$$

ou

$$\mathbf{x}(k+1) = \mathbf{G}\mathbf{x}(k) + \mathbf{H}u(k)$$

$$y(k) = \mathbf{C}\mathbf{x}(k) + Du(k)$$

onde

$$\mathbf{x}(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_{n-1}(k) \\ x_n(k) \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_{n-1} \\ h_n \end{bmatrix}$$

$$\mathbf{C} = [1 \quad 0 \quad \cdots \quad 0], \qquad D = h_0 = b_0$$

As condições iniciais $x_1(0), x_2(0), \ldots x_n(0)$ são determinadas de

$$x_1(0) = y(0) - h_0u(0)$$
$$x_2(0) = y(1) - h_0u(1) - h_1u(0)$$
$$x_3(0) = y(2) - h_0u(2) - h_1u(1) - h_2u(0)$$
$$\cdots$$
$$x_n(0) = y(n-1) - h_0u(n-1) - h_1u(n-2) - \cdots - h_{n-2}u(1) - h_{n-1}u(0)$$

Exemplo 14.11 Obtenha a representação de espaço de estados do sistema descrito por

$$y(k+2) + y(k+1) + 0{,}16y(k) = u(k+1) + 2u(k)$$

Definindo as variáveis de estados como segue,

$$x_1(k) = y(k)$$
$$x_2(k) = x_1(k+1) - u(k)$$

a equação de diferenças pode ser posta na representação padrão de espaço de estados:

$$x_1(k+1) = x_2(k) + u(k)$$
$$x_2(k+1) = -0{,}16x_1(k) - x_2(k) + u(k)$$
$$y(k) = x_1(k)$$

Reescrevendo

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -0{,}16 & -1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} [u(k)]$$

$$y(k) = [1 \quad 0] \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$$

As condições iniciais são dadas por

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} y(0) \\ y(1) - u(0) \end{bmatrix}$$

14.7 RESOLUÇÃO DA EQUAÇÃO DE ESTADO DE TEMPO DISCRETO

Nesta seção, inicialmente apresentaremos a solução da equação de estado de tempo discreto:

$$\mathbf{x}(k+1) = \mathbf{G}\mathbf{x}(k) + \mathbf{H}\mathbf{u}(k) \qquad (14.71)$$

usando um procedimento de recursão e, posteriormente, o método da transformada z. Então discutimos a discretização da equação de estado de tempo contínuo:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \qquad (14.72)$$

Isto é, derivamos a equação de estado de tempo discreto da forma dada pela Eq. (14.71) a partir da equação de estado de tempo contínuo, a Eq. (14.72).

Solução de equações de diferença. Em geral, equações de diferença são mais fáceis de resolver do que equações diferenciais porque podem ser resolvidas simplesmente através de um procedimento de recursão.

Como exemplo, considere a seguinte equação de diferenças:

$$x(k+1) + 0{,}2x(k) = 2u(k)$$

onde $x(0) = 0$ e $u(k) = 1$ para $k = 0, 1, 2, \ldots$. A solução $x(1)$ pode ser achada por recursão.

$$x(1) = -0{,}2x(0) + 2u(0) = 2$$

Similarmente,

$$x(2) = -0{,}2x(1) + 2u(1) = 1{,}6$$
$$x(3) = -0{,}2x(2) + 2u(2) = 1{,}68$$
$$x(4) = -0{,}2x(3) + 2u(3) = 1{,}664$$
$$\cdots$$

[Note que neste método $x(k+1)$ não pode ser computado a menos que $x(k)$ seja conhecido.] Este procedimento é bastante simples e conveniente para computação digital.

Solução de equações de estado de tempo discreto. O procedimento anterior de computar a solução de equações de diferenças escalares por recursão pode ser utilizado também para equação de diferenças matricial, ou a equação de estado de tempo discreto.

Considere a seguinte equação de estado e equação de saída:

$$\mathbf{x}(k+1) = \mathbf{G}\mathbf{x}(k) + \mathbf{H}\mathbf{u}(k) \qquad (14.73)$$
$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \qquad (14.74)$$

A solução da Eq. (14.73) para qualquer $k > 0$ pode ser obtida diretamente por recursão como segue:

$$\mathbf{x}(1) = \mathbf{G}\mathbf{x}(0) + \mathbf{H}\mathbf{u}(0)$$
$$\mathbf{x}(2) = \mathbf{G}\mathbf{x}(1) + \mathbf{H}\mathbf{u}(1) = \mathbf{G}^2\mathbf{x}(0) + \mathbf{G}\mathbf{H}\mathbf{u}(0) + \mathbf{H}\mathbf{u}(1)$$
$$\mathbf{x}(3) = \mathbf{G}\mathbf{x}(2) + \mathbf{H}\mathbf{u}(2) = \mathbf{G}^3\mathbf{x}(0) + \mathbf{G}^2\mathbf{H}\mathbf{u}(0) + \mathbf{G}\mathbf{H}\mathbf{u}(1) + \mathbf{H}\mathbf{u}(2)$$
$$\cdots$$

Repetindo este procedimento, obtemos

$$\mathbf{x}(k) = \mathbf{G}^k\mathbf{x}(0) + \sum_{j=0}^{k-1} \mathbf{G}^{k-j-1}\mathbf{H}\mathbf{u}(j) \qquad (k = 1, 2, 3, \ldots) \qquad (14.75)$$

Claramente $\mathbf{x}(k)$ consiste em duas partes, uma representando a contribuição do estado inicial $\mathbf{x}(0)$, e a outra a contribuição da entrada $\mathbf{u}(j)$, $j = 0, 1, 2, \ldots, k-1$.

Da Eq. (14.75), vemos que a matriz de transição de estado do sistema da Eq. (14.73) é

$$\mathbf{\Phi}(k) = \mathbf{G}^k \tag{14.76}$$

É uma matriz única satisfazendo

$$\mathbf{\Phi}(k+1) = \mathbf{G\Phi}(k), \qquad \mathbf{\Phi}(0) = \mathbf{I}$$

Em termos da matriz de transição de estados $\mathbf{\Phi}(k)$, a Eq. (14.75) pode ser escrita como

$$\mathbf{x}(k) = \mathbf{\Phi}(k)\mathbf{x}(0) + \sum_{j=0}^{k-1} \mathbf{\Phi}(k-j-1)\mathbf{Hu}(j) \tag{14.77}$$

$$= \mathbf{\Phi}(k)\mathbf{x}(0) + \sum_{j=0}^{k-1} \mathbf{\Phi}(j)\mathbf{Hu}(k-j-1) \tag{14.78}$$

Substituindo a Eq. (14.77) [ou Eq. (14.78)] na Eq. (14.74), a equação de saída pode ser escrita como

$$\mathbf{y}(k) = \mathbf{C\Phi}(k)\mathbf{x}(0) + \mathbf{C}\sum_{j=0}^{k-1} \mathbf{\Phi}(k-j-1)\mathbf{Hu}(j) + \mathbf{Du}(k)$$

$$= \mathbf{C\Phi}(k)\mathbf{x}(0) + \mathbf{C}\sum_{j=0}^{k-1} \mathbf{\Phi}(j)\mathbf{Hu}(k-j-1) + \mathbf{Du}(k)$$

Abordagem por transformada *z* para a solução de equações de estado de tempo discreto. Em seguida apresentamos a solução de equações de estado de tempo discreto usando o método da transformada z. Considere o sistema de tempo discreto descrito pela Eq. (14.73) reescrita como

$$\mathbf{x}(k+1) = \mathbf{Gx}(k) + \mathbf{Hu}(k) \tag{14.79}$$

Tomando a transformada z em ambos os lados da Eq. (14.79), obtemos

$$z\mathbf{X}(z) - z\mathbf{x}(0) = \mathbf{GX}(z) + \mathbf{HU}(z)$$

onde $\mathbf{X}(z) = \mathcal{Z}[\mathbf{x}(k)]$ e $\mathbf{U}(z) = \mathcal{Z}[\mathbf{u}(k)]$. Então,

$$(z\mathbf{I} - \mathbf{G})\mathbf{X}(z) = z\mathbf{x}(0) + \mathbf{HU}(z) \tag{14.80}$$

Pré-multiplicando ambos os lados da Eq. (14.80) por $(z\mathbf{I} - \mathbf{G})^{-1}$, obtemos

$$\mathbf{X}(z) = (z\mathbf{I} - \mathbf{G})^{-1}z\mathbf{x}(0) + (z\mathbf{I} - \mathbf{G})^{-1}\mathbf{HU}(z) \tag{14.81}$$

Tomando a transformada z em ambos os lados da Eq. (14.81), obtemos

$$\mathbf{x}(k) = \mathcal{Z}^{-1}[(z\mathbf{I} - \mathbf{G})^{-1}z]\mathbf{x}(0) + \mathcal{Z}^{-1}[(z\mathbf{I} - \mathbf{G})^{-1}\mathbf{HU}(z)] \tag{14.82}$$

Comparando as Eqs. (14.75) e (14.82), obtemos

$$\mathbf{G}^k = \mathcal{Z}^{-1}[(z\mathbf{I} - \mathbf{G})^{-1}z] \tag{14.83}$$

$$\sum_{j=0}^{k-1} \mathbf{G}^{k-j-1}\mathbf{Hu}(j) = \mathcal{Z}^{-1}[(z\mathbf{I} - \mathbf{G})^{-1}\mathbf{HU}(z)] \tag{14.84}$$

onde $k = 1, 2, 3, \ldots$. [Para a Eq. (14.83), veja o Problema A.13.1.]

A Eq. (14.84) pode também ser obtida diretamente. Tomando a transformada z de

$$\sum_{j=0}^{k-1} \mathbf{G}^{k-j-1}\mathbf{Hu}(j) \quad \text{onde} \quad k = 1, 2, 3, \ldots, \text{ obtemos}$$

$$\mathcal{Z}\Big[\sum_{j=0}^{k-1} \mathbf{G}^{k-j-1}\mathbf{Hu}(j)\Big] = \sum_{k=0}^{\infty}\sum_{j=0}^{k-1} \mathbf{G}^{k-j-1}\mathbf{Hu}(j)z^{-k}$$

$$= \sum_{k=0}^{\infty} z^{-k+j}\sum_{j=0}^{k-1} \mathbf{G}^{k-j-1}\mathbf{Hu}(j)z^{-j}$$

$$= \sum_{k=1}^{\infty} z^{-k+j}\sum_{j=0}^{k-1} \mathbf{G}^{k-j-1}\mathbf{Hu}(j)z^{-j}$$

$$= \sum_{k=1}^{\infty} [\mathbf{G}^{k-1}\mathbf{Hu}(0)z^{-k} + \mathbf{G}^{k-2}\mathbf{Hu}(1)z^{-k} + \mathbf{G}^{k-3}\mathbf{Hu}(2)z^{-k} + \cdots]$$

$$= (\mathbf{H}z^{-1} + \mathbf{GH}z^{-2} + \mathbf{G}^2\mathbf{H}z^{-3} + \cdots) \times [\mathbf{u}(0) + \mathbf{u}(1)z^{-1} + \mathbf{u}(2)z^{-2} + \cdots]$$

$$= (\mathbf{I} + \mathbf{G}z^{-1} + \mathbf{G}^2z^{-2} + \cdots)\mathbf{H}z^{-1} \times [\mathbf{u}(0) + \mathbf{u}(1)z^{-1} + \mathbf{u}(2)z^{-2} + \cdots]$$

$$= (\mathbf{I} - \mathbf{G}z^{-1})^{-1}\mathbf{H}z^{-1}\sum_{k=0}^{\infty} \mathbf{u}(k)z^{-k}$$

$$= (z\mathbf{I} - \mathbf{G})^{-1}\mathbf{HU}(z)$$

Finalmente, note que da Eq. (14.81) podemos ver que a equação característica do sistema de tempo discreto é

$$|z\mathbf{I} - \mathbf{G}| = 0 \tag{14.85}$$

Referindo-se à Seção 13.6, sabemos que o sistema de tempo discreto é estável se e apenas se todas as raízes da equação característica, a Eq. (14.85), estão no círculo unitário centrado na origem do plano z.

Exemplo 14.12 Obtenha a matriz de transição de estado do seguinte sistema de tempo discreto:

$$\mathbf{x}(k+1) = \mathbf{Gx}(k) + \mathbf{H}u(k)$$

onde

$$\mathbf{G} = \begin{bmatrix} 0 & 1 \\ -0{,}16 & -1 \end{bmatrix}, \qquad \mathbf{H} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Em seguida obtenha $\mathbf{x}(k)$ quando $u(k) = 1$ para $k = 0, 1, 2, \ldots$. Suponha que a condição inicial é dada por

$$\mathbf{x}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Das Eqs. (14.76) e (14.83), vemos que a matriz de transição de estado $\mathbf{\Phi}(k)$ é

$$\mathbf{\Phi}(k) = \mathbf{G}^k = \mathcal{Z}^{-1}[(z\mathbf{I} - \mathbf{G})^{-1}z]$$

Portanto, primeiro obtemos $(z\mathbf{I} - \mathbf{G})^{-1}$.

$$(z\mathbf{I} - \mathbf{G})^{-1} = \begin{bmatrix} z & -1 \\ 0{,}16 & z+1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \dfrac{z+1}{(z+0{,}2)(z+0{,}8)} & \dfrac{1}{(z+0{,}2)(z+0{,}8)} \\ \dfrac{-0{,}16}{(z+0{,}2)(z+0{,}8)} & \dfrac{z}{(z+0{,}2)(z+0{,}8)} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\frac{4}{3}}{z+0{,}2} + \dfrac{-\frac{1}{3}}{z+0{,}8} & \dfrac{\frac{5}{3}}{z+0{,}2} + \dfrac{-\frac{5}{3}}{z+0{,}8} \\ \dfrac{-\frac{0{,}8}{3}}{z+0{,}2} + \dfrac{\frac{0{,}8}{3}}{z+0{,}8} & \dfrac{-\frac{1}{3}}{z+0{,}2} + \dfrac{\frac{4}{3}}{z+0{,}8} \end{bmatrix}$$

Portanto $\mathbf{\Phi}(k)$ é obtido como

$$\mathbf{\Phi}(k) = \mathbf{G}^k = \mathcal{Z}^{-1}[(z\mathbf{I} - \mathbf{G})^{-1}z]$$

$$= \mathcal{Z}^{-1}\begin{bmatrix} \frac{4}{3}\left(\frac{z}{z+0{,}2}\right) - \frac{1}{3}\left(\frac{z}{z+0{,}8}\right) & \frac{5}{3}\left(\frac{z}{z+0{,}2}\right) - \frac{5}{3}\left(\frac{z}{z+0{,}8}\right) \\ -\frac{0{,}8}{3}\left(\frac{z}{z+0{,}2}\right) + \frac{0{,}8}{3}\left(\frac{z}{z+0{,}8}\right) & -\frac{1}{3}\left(\frac{z}{z+0{,}2}\right) + \frac{4}{3}\left(\frac{z}{z+0{,}8}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{4}{3}(-0{,}2)^k - \frac{1}{3}(-0{,}8)^k & \frac{5}{3}(-0{,}2)^k - \frac{5}{3}(-0{,}8)^k \\ -\frac{0{,}8}{3}(-0{,}2)^k + \frac{0{,}8}{3}(-0{,}8)^k & -\frac{1}{3}(-0{,}2)^k + \frac{4}{3}(-0{,}8)^k \end{bmatrix}$$

Em seguida, calcule $\mathbf{x}(k)$. A transformada z de $\mathbf{x}(k)$ é dada por

$$\mathcal{Z}[\mathbf{x}(k)] = \mathbf{X}(z) = (z\mathbf{I} - \mathbf{G})^{-1}z\mathbf{x}(0) + (z\mathbf{I} - \mathbf{G})^{-1}\mathbf{H}U(z)$$
$$= (z\mathbf{I} - \mathbf{G})^{-1}[z\mathbf{x}(0) + \mathbf{H}U(z)]$$

Como

$$U(z) = \frac{z}{z-1}$$

obtemos

$$z\mathbf{x}(0) + \mathbf{H}U(z) = \begin{bmatrix} z \\ -z \end{bmatrix} + \begin{bmatrix} \dfrac{z}{z-1} \\ \dfrac{z}{z-1} \end{bmatrix} = \begin{bmatrix} \dfrac{z^2}{z-1} \\ \dfrac{-z^2+2z}{z-1} \end{bmatrix}$$

Portanto

$$\mathbf{X}(z) = (z\mathbf{I} - \mathbf{G})^{-1}[z\mathbf{x}(0) + \mathbf{H}U(z)]$$

$$= \begin{bmatrix} \dfrac{(z^2+2)z}{(z+0{,}2)(z+0{,}8)(z-1)} \\ \dfrac{(-z^2+1{,}84z)z}{(z+0{,}2)(z+0{,}8)(z-1)} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{-\frac{17}{6}z}{z+0{,}2} + \dfrac{\frac{22}{9}z}{z+0{,}8} + \dfrac{\frac{25}{18}z}{z-1} \\ \dfrac{\frac{3{,}4}{6}z}{z+0{,}2} + \dfrac{-\frac{17{,}6}{9}z}{z+0{,}8} + \dfrac{\frac{7}{18}z}{z-1} \end{bmatrix}$$

Portanto

$$\mathbf{x}(k) = \mathcal{Z}^{-1}[\mathbf{X}(z)] = \begin{bmatrix} -\frac{17}{6}(-0{,}2)^k + \frac{22}{9}(-0{,}8)^k + \frac{25}{18} \\ \frac{3{,}4}{6}(-0{,}2)^k - \frac{17{,}6}{9}(-0{,}8)^k + \frac{7}{18} \end{bmatrix}$$

Discretização de equações de estado de tempo contínuo. Se desejamos computar o estado $\mathbf{x}(t)$ usando um computador digital, devemos converter uma equação de estado de tempo contínuo para uma equação de estado de tempo discreto. Em seguida, apresentaremos tal procedimento. Supomos que o vetor de entrada $\mathbf{u}(t)$ muda apenas nos instantes de amostragem igualmente espaçados. (A operação de amostragem feita aqui é fictícia. Derivaremos a equação de estado de tempo discreto que fornece valores exatos em $t = kT$, $k = 0, 1, 2, \ldots$.)

Considere a equação de estado de tempo contínuo

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{14.86}$$

A seguir, para tornar a análise mais clara, usaremos a notação kT e $(k+1)T$ em vez de k e $k+1$. A representação de tempo discreto da Eq. (14.86) tomará a forma

$$\mathbf{x}((k+1)T) = \mathbf{G}(T)\mathbf{x}(kT) + \mathbf{H}(T)\mathbf{u}(kT) \tag{14.87}$$

Note que as matrizes $\mathbf{G}$ e $\mathbf{H}$ dependem do período de amostragem T. (Uma vez que este período é fixado, $\mathbf{G}$ e $\mathbf{H}$ são matrizes constantes.)

Para determinar $\mathbf{G}(T)$ e $\mathbf{H}(T)$, usamos a solução da Eq. (14.86), ou

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + e^{\mathbf{A}t}\int_0^t e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)\,d\tau$$

Supomos que todos os componentes de $\mathbf{u}(t)$ são constantes no intervalo entre

quaisquer dois instantes sucessivos de amostragem, ou $\mathbf{u}(t) = \mathbf{u}(kT)$ para o k-ésimo período de amostragem. Como

$$\mathbf{x}((k+1)T) = e^{\mathbf{A}(k+1)T}\mathbf{x}(0) + e^{\mathbf{A}(k+1)T}\int_0^{(k+1)T} e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)\,d\tau \qquad (14.88)$$

e

$$\mathbf{x}(kT) = e^{\mathbf{A}kT}\mathbf{x}(0) + e^{\mathbf{A}kT}\int_0^{kT} e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)\,d\tau \qquad (14.89)$$

multiplicando a Eq. (14.89) por $e^{\mathbf{A}T}$ e subtraindo da Eq. (14.88), obtemos

$$\begin{aligned}\mathbf{x}((k+1)T) &= e^{\mathbf{A}T}\mathbf{x}(kT) + e^{\mathbf{A}(k+1)T}\int_{kT}^{(k+1)T} e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)\,d\tau \\ &= e^{\mathbf{A}T}\mathbf{x}(kT) + e^{\mathbf{A}T}\int_0^{T} e^{-\mathbf{A}t}\mathbf{B}\mathbf{u}(kT)\,dt \\ &= e^{\mathbf{A}T}\mathbf{x}(kT) + \int_0^{T} e^{\mathbf{A}\lambda}\mathbf{B}\mathbf{u}(kT)\,d\lambda \end{aligned} \qquad (14.90)$$

onde $\lambda = T - t$. Se definirmos

$$\mathbf{G}(T) = e^{\mathbf{A}T} \qquad (14.91)$$

$$\mathbf{H}(T) = \left(\int_0^T e^{\mathbf{A}t}\,dt\right)\mathbf{B} \qquad (14.92)$$

então a Eq. (14.90) se torna

$$\mathbf{x}((k+1)T) = \mathbf{G}(T)\mathbf{x}(kT) + \mathbf{H}(T)\,\mathbf{u}(kT)$$

que é a Eq. (14.87). Portanto, as Eqs. (14.91) e (14.92) fornecem as matrizes $\mathbf{G}(T)$ e $\mathbf{H}(T)$ desejadas.

Exemplo 14.13 Obtenha uma representação por espaço de estados de tempo discreto do seguinte sistema de tempo contínuo:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix}[u]$$

A equação de estado de tempo discreto desejada será da seguinte forma:

$$\mathbf{x}((k+1)T) = \mathbf{G}(T)\mathbf{x}(kT) + \mathbf{H}(T)u(kT)$$

As matrizes $\mathbf{G}(T)$ e $\mathbf{H}(T)$ podem ser obtidas das Eqs. (14.91) e (14.92) como

$$\mathbf{G}(T) = e^{\mathbf{A}T} = \begin{bmatrix}1 & \frac{1}{2}(1 - e^{-2T})\\ 0 & e^{-2T}\end{bmatrix}$$

$$\begin{aligned}\mathbf{H}(T) &= \left(\int_0^T e^{\mathbf{A}t}\,dt\right)\mathbf{B} \\ &= \left\{\int_0^T \begin{bmatrix}1 & \frac{1}{2}(1 - e^{-2t})\\ 0 & e^{-2t}\end{bmatrix} dt\right\}\begin{bmatrix}0\\ 1\end{bmatrix} \\ &= \begin{bmatrix}\frac{1}{2}\left(T + \frac{e^{-2T} - 1}{2}\right)\\ \frac{1}{2}(1 - e^{-2T})\end{bmatrix}\end{aligned}$$

Portanto

$$\begin{bmatrix}x_1((k+1)T)\\ x_2((k+1)T)\end{bmatrix} = \begin{bmatrix}1 & \frac{1}{2}(1 - e^{-2T})\\ 0 & e^{-2T}\end{bmatrix}\begin{bmatrix}x_1(kT)\\ x_2(kT)\end{bmatrix} + \begin{bmatrix}\frac{1}{2}\left(T + \frac{e^{-2T} - 1}{2}\right)\\ \frac{1}{2}(1 - e^{-2T})\end{bmatrix}u(kT)$$

Se, por exemplo, o período de amostragem é de 1 segundo, ou $T = 1$, então a equação de estado de tempo discreto se torna

$$\begin{bmatrix}x_1(k+1)\\ x_2(k+1)\end{bmatrix} = \begin{bmatrix}1 & 0{,}432\\ 0 & 0{,}135\end{bmatrix}\begin{bmatrix}x_1(k)\\ x_2(k)\end{bmatrix} + \begin{bmatrix}0{,}284\\ 0{,}432\end{bmatrix}[u(k)]$$

Problemas Ilustrativos e Soluções

Problema A.14.1 Obtenha a representação de espaço de estados do sistema visto na Fig. 14.11.

Solução. Neste problema, ilustraremos um método de derivação de uma representação por espaço de estados de sistemas dados na forma de diagrama de blocos.

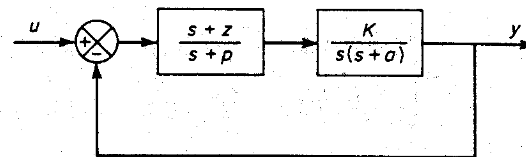


Fig. 14.11 Sistema de controle.

No presente problema, inicialmente expandiremos $(s + z)/(s + p)$ em frações parciais.

$$\frac{s+z}{s+p} = 1 + \frac{z-p}{s+p}$$

Em seguida convertemos $K/[s(s + a)]$ no produto de K/s e $1/(s + a)$. Em seguida redesenhamos o diagrama de blocos, como visto na Fig. 14.12. Definindo um conjunto de variáveis de

estado, como visto na Fig. 14.12, obtemos as seguintes equações:

$$\dot{x}_1 = -ax_1 + x_2$$
$$\dot{x}_2 = -Kx_1 + Kx_3 + Ku$$
$$\dot{x}_3 = -(z-p)x_1 - px_3 + (z-p)u$$
$$y = x_1$$

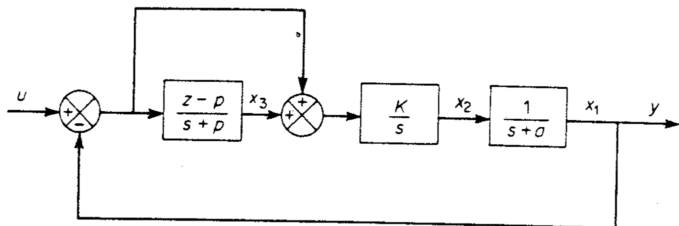


Fig. 14.12 Diagrama de blocos definindo as variáveis de estado para o sistema visto na Fig. 14.11.

Reescrevendo, obtemos

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -a & 1 & 0 \\ -K & 0 & K \\ -(z-p) & 0 & -p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ K \\ z-p \end{bmatrix} [u]$$

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Problema A.14.2 Derive uma equação de estado para o sistema visto na Fig. 14.13. Seja x_1 a corrente pela indutância L, ou seja, x_1 $i_1 - i_2$ e a tensão no capacitor x_2. Suponha que $e(t)$ é a entrada do sistema. Considere x_1 e x_2 como as variáveis de estado para o sistema.

Solução. Da Fig. 14.13, obtemos as seguintes equações:

$$L\left(\frac{di_1}{dt} - \frac{di_2}{dt}\right) + R_1 i_1 = e(t)$$

$$\frac{1}{C}\int i_2 dt + R_2 i_2 + L\left(\frac{di_2}{dt} - \frac{di_1}{dt}\right) = 0$$

Em termos de x_1 e x_2, podemos reescrever estas equações como

$$L\frac{dx_1}{dt} + R_1(x_1 + i_2) = e(t)$$

$$x_2 + R_2 i_2 + R_1(x_1 + i_2) = e(t)$$

$$x_2 = \frac{1}{C}\int i_2\, dt$$

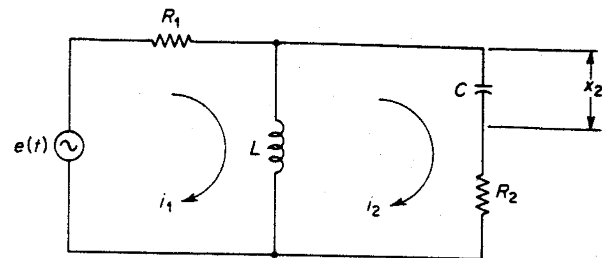


Fig. 14.13 Sistema composto por uma rede elétrica.

Para obter a equação de estado, devemos eliminar i_2. Para tal, tomemos a transformada de Laplace destas equações:

$$L[sX_1(s) - x_1(0)] + R_1[X_1(s) + I_2(s)] = E(s)$$

$$X_2(s) + R_2 I_2(s) + R_1[X_1(s) + I_2(s)] = E(s)$$

$$sX_2(s) - x_2(0) = \frac{1}{C}I_2(s)$$

Eliminando $I_2(s)$ das últimas três equações, temos

$$L[sX_1(s) - x_1(0)] = \frac{-R_1R_2}{R_1+R_2}X_1(s) + \frac{R_1}{R_1+R_2}X_2(s) + \frac{R_2}{R_1+R_2}E(s)$$

$$sX_2(s) - x_2(0) = \frac{-R_1}{C(R_1+R_2)}X_1(s) + \frac{-1}{C(R_1+R_2)}X_2(s) + \frac{1}{C(R_1+R_2)}E(s)$$

Tomando a transformada inversa de Laplace, obtém-se

$$\dot{x}_1(t) = \frac{-R_1R_2}{L(R_1+R_2)}x_1(t) + \frac{R_1}{L(R_1+R_2)}x_2(t) + \frac{R_2}{L(R_1+R_2)}e(t)$$

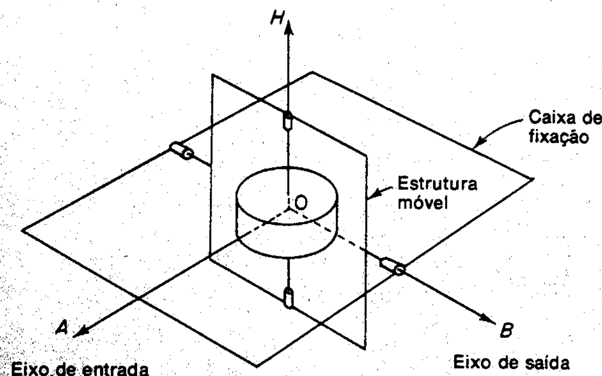
$$\dot{x}_2(t) = \frac{-R_1}{C(R_1+R_2)}x_1(t) + \frac{-1}{C(R_1+R_2)}x_2(t) + \frac{1}{C(R_1+R_2)}e(t)$$

A equação de estado para o sistema é então dada por

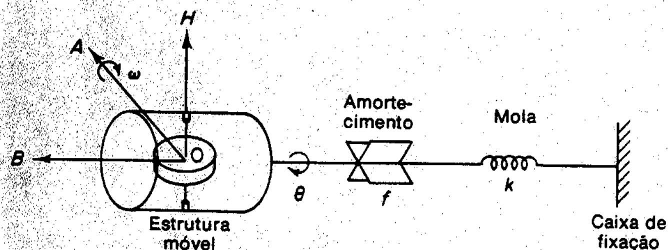
$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} \frac{-R_1R_2}{L(R_1+R_2)} & \frac{R_1}{L(R_1+R_2)} \\ \frac{-R_1}{C(R_1+R_2)} & \frac{-1}{C(R_1+R_2)} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} \frac{R_2}{L(R_1+R_2)} \\ \frac{1}{C(R_1+R_2)} \end{bmatrix} [e(t)]$$

Problema A.14.3 Giroscópios para medir movimento angular são comumente empregados em sistemas de navegação inercial, sistemas de autopiloto etc.

A Fig. 14.14(a) mostra um giroscópio com um único grau de liberdade. A roda que gira está montada em uma estrutura móvel que, por sua vez, está montada na caixa do giroscópio. A estrutura móvel pode-se mover livremente em relação à caixa, em torno do eixo de saída OB. Note que o eixo de saída é perpendicular ao eixo da roda que gira. O eixo de entrada em torno do qual uma velocidade de rotação, ou ângulo, é medida é perpendicular tanto ao eixo de



(a)



(b)

Fig. 14.14 (a) Diagrama esquemático de um giroscópio com um grau de liberdade; (b) diagrama funcional do giroscópio visto em (a).

saída como de rotação da roda. A informação sobre o sinal de entrada (a velocidade de rotação ou ângulo em torno do eixo de entrada) é obtida a partir do movimento resultante da estrutura móvel em torno do eixo de saída, relativamente à caixa.

A Fig. 14.14(b) mostra um diagrama funcional do sistema de giroscópio. A equação do movimento em torno do eixo de saída pode ser obtida equacionando-se a taxa de variação da quantidade de movimento angular com a soma dos torques externos.

A variação de quantidade de movimento angular em torno do eixo OB consiste em duas partes: $I\ddot{\theta}$, a variação devida à aceleração da estrutura móvel em torno do eixo OB, e $-H\omega\cos\theta$, a variação devido ao vetor de quantidade de movimento angular da roda em rotação em torno do eixo OA. Os torques externos consistem em $-f\dot{\theta}$, o torque de amortecimento, e $-k\theta$, o torque da mola. Portanto a equação do sistema de giroscópio é

$$I\ddot{\theta} - H\omega\cos\theta = -f\dot{\theta} - k\theta$$

ou

$$I\ddot{\theta} + f\dot{\theta} + k\theta = H\omega\cos\theta \qquad (14.93)$$

Na prática, θ é um ângulo muito pequeno, normalmente não maior do que $\pm$ 2,5 graus.

Obtenha a representação de espaço de estados do sistema de giroscópio.

Solução. Neste sistema, θ e $\dot{\theta}$ podem ser escolhidos como variáveis de estado. A variável de entrada é ω. Vamos definir

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}, \qquad u = \omega$$

Então a Eq. (14.93) pode ser escrita como segue:

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -\frac{k}{I}x_1 - \frac{f}{I}x_2 + \frac{H}{I}u\cos x_1$$

ou

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u)$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad \mathbf{f}(\mathbf{x}, u) = \begin{bmatrix} f_1(\mathbf{x}, u) \\ f_2(\mathbf{x}, u) \end{bmatrix} = \begin{bmatrix} x_2 \\ -\frac{k}{I}x_1 - \frac{f}{I}x_2 + \frac{H}{I}u\cos x_1 \end{bmatrix}$$

Claramente $f_2(\mathbf{x}, u)$ envolve um termo não linear em x_1 e u. Expandindo $\cos x_1$ nesta representação em série,

$$\cos x_1 = 1 - \frac{1}{2}x_1^2 + \cdots$$

e observando que x_1 é um ângulo muito pequeno, podemos aproximar $\cos x_1$ como sendo 1, obtendo a seguinte equação de estado linearizada:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{I} & -\frac{f}{I} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{H}{I} \end{bmatrix} [u]$$

Se θ é considerado como a saída y, então

$$y = \theta$$

ou

$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Problema A.14.4 Considere o sistema descrito por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

onde

$\mathbf{x}$ = vetor n-dimensional
$\mathbf{u}$ = vetor r-dimensional
$\mathbf{A}$ = matriz constante $n \times n$
$\mathbf{B}$ = matriz constante $n \times r$

Obtenha a resposta para cada uma das seguintes entradas:

1. Os r componentes de $\mathbf{u}$ são funções impulsivas de várias amplitudes.
2. Os r componentes de $\mathbf{u}$ são funções-degrau de várias amplitudes.
3. Os r componentes de $\mathbf{u}$ são funções-rampa de várias amplitudes.

Suponha que cada entrada é aplicada em $t = 0$.

Solução. Substituindo $t_0 = 0-$ na Eq. (14.47), podemos escrever a solução $\mathbf{x}(t)$ como segue:

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0-) + \int_{0-}^{t} e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau$$

1. *Resposta ao impulso:* Vamos escrever a entrada impulsiva $\mathbf{u}(t)$ como

$$\mathbf{u}(t) = \delta(t)\mathbf{w}$$

onde $\mathbf{w}$ é um vetor cujos componentes são as amplitudes das r funções impulsivas aplicadas em $t = 0$. A solução ao impulso dado em $t = 0$ é

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0-) + \int_{0-}^{t} e^{\mathbf{A}(t-\tau)}\mathbf{B}\delta(\tau)\mathbf{w}\,d\tau$$
$$= e^{\mathbf{A}t}\mathbf{x}(0-) + e^{\mathbf{A}t}\mathbf{B}\mathbf{w}$$

2. *Resposta ao degrau:* Vamos escrever a entrada em degrau $\mathbf{u}(t)$ como

$$\mathbf{u}(t) = \mathbf{k}$$

onde $\mathbf{k}$ é um vetor cujos componentes são as amplitudes de r funções-degrau aplicadas em $t = 0$. A solução para a entrada em degrau em $t = 0$ é dada por

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_{0}^{t} e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{k}\,d\tau$$
$$= e^{\mathbf{A}t}\mathbf{x}(0) + e^{\mathbf{A}t}\left[\int_0^t \left(\mathbf{I} - \mathbf{A}\tau + \frac{\mathbf{A}^2\tau^2}{2!} - \cdots\right)d\tau\right]\mathbf{B}\mathbf{k}$$
$$= e^{\mathbf{A}t}\mathbf{x}(0) + e^{\mathbf{A}t}\left(\mathbf{I}t - \frac{\mathbf{A}t^2}{2!} + \frac{\mathbf{A}^2t^3}{3!} - \cdots\right)\mathbf{B}\mathbf{k}$$

Se $\mathbf{A}$ é não singular, então esta última equação pode ser simplificada para resultar em

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + e^{\mathbf{A}t}[-(\mathbf{A}^{-1})(e^{-\mathbf{A}t} - \mathbf{I})]\mathbf{B}\mathbf{k}$$
$$= e^{\mathbf{A}t}\mathbf{x}(0) + \mathbf{A}^{-1}(e^{\mathbf{A}t} - \mathbf{I})\mathbf{B}\mathbf{k}$$

3. *Resposta à rampa:* Vamos escrever a entrada em rampa $\mathbf{u}(t)$ como

$$\mathbf{u}(t) = t\mathbf{v}$$

onde $\mathbf{v}$ é um vetor cujos componentes são as amplitudes das funções em rampa aplicada em $t = 0$. A solução para uma entrada em rampa $t\mathbf{v}$ aplicada em $t = 0$ é

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_{0}^{t} e^{\mathbf{A}(t-\tau)}\mathbf{B}\tau\mathbf{v}\,d\tau$$
$$= e^{\mathbf{A}t}\mathbf{x}(0) + e^{\mathbf{A}t}\int_{0}^{t} e^{-\mathbf{A}\tau}\tau\,d\tau\,\mathbf{B}\mathbf{v}$$
$$= e^{\mathbf{A}t}\mathbf{x}(0) + e^{\mathbf{A}t}\left(\frac{\mathbf{I}}{2}t^2 - \frac{2\mathbf{A}}{3!}t^3 + \frac{3\mathbf{A}^2}{4!}t^4 - \frac{4\mathbf{A}^3}{5!}t^5 + \cdots\right)\mathbf{B}\mathbf{v}$$

Se $\mathbf{A}$ é não singular, então esta última equação pode ser simplificada para fornecer

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + (\mathbf{A}^{-2})(e^{\mathbf{A}t} - \mathbf{I} - \mathbf{A}t)\mathbf{B}\mathbf{v}$$
$$= e^{\mathbf{A}t}\mathbf{x}(0) + [\mathbf{A}^{-2}(e^{\mathbf{A}t} - \mathbf{I}) - \mathbf{A}^{-1}t]\mathbf{B}\mathbf{v}$$

Problema A.14.5 Um método útil para calcular $e^{\mathbf{A}t}$ é baseado no teorema de Cayley-Hamilton, que diz que uma matriz $\mathbf{A}$ $n \times n$ satisfaz sua própria equação característica. Isto permite escrever todas as potências de $\mathbf{A}$ em termos de um polinômio em $\mathbf{A}$ de grau igual ou menor do que $n - 1$, isto é,

$$e^{\mathbf{A}t} = \alpha_0(t)\mathbf{I} + \alpha_1(t)\mathbf{A} + \cdots + \alpha_{n-1}(t)\mathbf{A}^{n-1} \qquad (14.94)$$

Se $\mathbf{A}$ tem autovalores distintos, então, para determinar os coeficientes $\alpha_0(t), \alpha_1(t), \ldots, \alpha_{n-1}(t)$, substituímos cada um dos autovalores de $\mathbf{A}$ na versão escalar da Eq. (14.94) obtendo n equações simultâneas. (Note que se os autovalores de $\mathbf{A}$ são múltiplos, então o procedimento deve ser levemente modificado. Para detalhes, referir-se ao Cap. 6 da Referência O-1.)

Usando a presente abordagem, obtenha $e^{\mathbf{A}t}$, onde

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$$

Solução. Os autovalores de $\mathbf{A}$ são obtidos de

$$|\lambda\mathbf{I} - \mathbf{A}| = 0$$

ou

$$\begin{vmatrix} \lambda & -1 \\ 2 & \lambda + 3 \end{vmatrix} = 0$$

Portanto

$$\lambda_1 = -1, \qquad \lambda_2 = -2$$

$e^{\mathbf{A}t}$ pode ser expressa como

$$e^{\mathbf{A}t} = \alpha_0(t)\mathbf{I} + \alpha_1(t)\mathbf{A} \qquad (14.95)$$

A versão escalar desta equação é

$$e^{\lambda t} = \alpha_0(t) + \alpha_1(t)\lambda \qquad (14.96)$$

Substituindo $\lambda = \lambda_1 = -1$ na Eq. (14.96), obtém-se

$$e^{-t} = \alpha_0(t) - \alpha_1(t) \qquad (14.97)$$

De forma semelhante, substituindo $\lambda = \lambda_2 = -2$ na Eq. (14.96), temos

$$e^{-2t} = \alpha_0(t) - 2\alpha_1(t) \tag{14.98}$$

Resolvendo as Eqs. (14.97) e (14.98) para obter $\alpha_0(t)$ e $\alpha_1(t)$, obtemos

$$\alpha_0(t) = 2e^{-t} - e^{-2t}$$
$$\alpha_1(t) = e^{-t} - e^{-2t}$$

Substituindo estes valores na Eq. (14.95), resulta

$$e^{\mathbf{A}t} = (2e^{-t} - e^{-2t})\mathbf{I} + (e^{-t} - e^{-2t})\mathbf{A} = \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix}$$

Problema A.14.6 Considere o sistema descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad a_{21} \neq 0$$

Supondo que os dois autovalores λ_1 e λ_2 da matriz de coeficientes são distintos, determine uma matriz de transformação $\mathbf{P}$ que transforme a primeira matriz em diagonal. Ache $\mathbf{P}$ em função de a_{ij}, λ_1 e λ_2.

Solução.

$$\mathbf{P} = \begin{bmatrix} a_{22} - \lambda_1 & a_{22} - \lambda_2 \\ -a_{21} & -a_{21} \end{bmatrix}$$

ou

$$\mathbf{P} = \begin{bmatrix} \lambda_2 - a_{11} & \lambda_1 - a_{11} \\ -a_{21} & -a_{21} \end{bmatrix}$$

Problema A.14.7 Obtenha uma representação de espaço de estados de tempo discreto do sistema visto na Fig. 14.15.

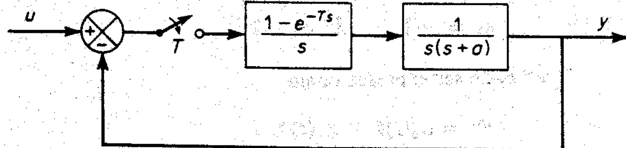


Fig. 14.15 Sistema de tempo discreto.

Solução. A função de transferência de malha-aberta do sistema é

$$G(s) = \frac{1 - e^{-Ts}}{s} \frac{1}{s(s + a)}$$

A transformada z de $G(s)$ é dada por

$$G(z) = \frac{T}{a(z-1)} - \frac{1 - e^{-aT}}{a^2(z - e^{-aT})}$$

Em termos da função de transferência pulsada, a Fig. 14.15 pode ser modificada como visto na Fig. 14.16.

Vamos definir as variáveis de estado como visto na Fig. 14.16. Observando que $z\mathcal{Z}[\mathbf{x}(k)] = \mathcal{Z}[\mathbf{x}(k+1)]$, obtemos a seguinte equação de estado de tempo discreto:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 1 - \frac{T}{a} & -\frac{T}{a} \\ \frac{1 - e^{-aT}}{a^2} & \frac{1 + (a^2 - 1)e^{-aT}}{a^2} \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} \frac{T}{a} \\ \frac{-1 + a^{-aT}}{a^2} \end{bmatrix} [u(k)]$$

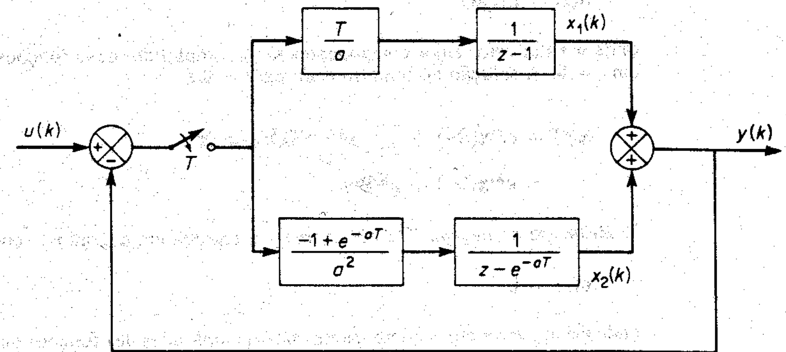


Fig. 14.16 Diagrama de blocos definindo as variáveis de estado para o sistema visto na Fig. 14.15.

Problema A.14.8 O país A tem uma população de 100 milhões em 1970. A população da cidade B no país A tem 10 milhões em 1970.

Suponha que a cada ano 4% da população do ano anterior da cidade B abandonam esta cidade, e que 2% da população do ano anterior do país fora da cidade B mudam para a cidade B.

Determine a população da cidade B no ano de 1980. Suponha que o aumento natural da população é de 1% ao ano.

Solução. Vamos chamar o ano 1970 de ano 0. Seja a população da cidade B no k-ésimo ano $x_1(k)$ e a população do país A, excetuando-se a cidade B, $x_2(k)$. Então obtemos as seguintes equações:

$$x_1(k+1) = 1{,}01[(1 - 0{,}04)x_1(k) + 0{,}02x_2(k)]$$
$$x_2(k+1) = 1{,}01[0{,}04x_1(k) + (1 - 0{,}02)x_2(k)]$$

com as condições iniciais

$$x_1(0) = 10 \times 10^6, \qquad x_2(0) = 90 \times 10^6$$

Em termos de notação matricial,

$$\mathbf{x}(k+1) = \mathbf{G}\mathbf{x}(k) \tag{14.99}$$

onde

$$\mathbf{G} = \begin{bmatrix} (1{,}01)(1-0{,}04) & (1{,}01)(0{,}02) \\ (1{,}01)(0{,}04) & (1{,}01)(1-0{,}02) \end{bmatrix}$$

A solução da Eq. (14.99) é dada por

$$\mathbf{x}(k) = \mathbf{G}^k\mathbf{x}(0)$$

Referindo-nos à Eq. (14.83), temos

$$\mathbf{G}^k = \mathcal{Z}^{-1}[(z\mathbf{I}-\mathbf{G})^{-1}z]$$

Como

$$(z\mathbf{I}-\mathbf{G})^{-1} = \begin{bmatrix} z-(1{,}01)(0{,}96) & -(1{,}01)(0{,}02) \\ -(1{,}01)(0{,}04) & z-(1{,}01)(0{,}98) \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \dfrac{z-(1{,}01)(0{,}98)}{(z-1{,}01)(z-0{,}9494)} & \dfrac{(1{,}01)(0{,}02)}{(z-1{,}01)(z-0{,}9494)} \\ \dfrac{(1{,}01)(0{,}04)}{(z-1{,}01)(z-0{,}9494)} & \dfrac{z-(1{,}01)(0{,}96)}{(z-1{,}01)(z-0{,}9494)} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\frac{1}{3}}{z-1{,}01} + \dfrac{\frac{2}{3}}{z-0{,}9494} & \dfrac{\frac{1}{3}}{z-1{,}01} + \dfrac{-\frac{1}{3}}{z-0{,}9494} \\ \dfrac{\frac{2}{3}}{z-1{,}01} + \dfrac{-\frac{2}{3}}{z-0{,}9494} & \dfrac{\frac{2}{3}}{z-1{,}01} + \dfrac{\frac{1}{3}}{z-0{,}9494} \end{bmatrix}$$

$\mathbf{G}^k$ pode ser obtido como

$$\mathbf{G}^k = \mathcal{Z}^{-1}[(z\mathbf{I}-\mathbf{G})^{-1}z]$$

$$= \mathcal{Z}^{-1}\begin{bmatrix} \frac{1}{3}\dfrac{z}{z-1{,}01} + \frac{2}{3}\dfrac{z}{z-0{,}9494} & \frac{1}{3}\dfrac{z}{z-1{,}01} - \frac{1}{3}\dfrac{z}{z-0{,}9494} \\ \frac{2}{3}\dfrac{z}{z-1{,}01} - \frac{2}{3}\dfrac{z}{z-0{,}9494} & \frac{2}{3}\dfrac{z}{z-1{,}01} + \frac{1}{3}\dfrac{z}{z-0{,}9494} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{3}(1{,}01)^k + \frac{2}{3}(0{,}9494)^k & \frac{1}{3}(1{,}01)^k - \frac{1}{3}(0{,}9494)^k \\ \frac{2}{3}(1{,}01)^k - \frac{2}{3}(0{,}9494)^k & \frac{2}{3}(1{,}01)^k + \frac{1}{3}(0{,}9494)^k \end{bmatrix}$$

Portanto $\mathbf{x}(k)$ é obtido como

$$\mathbf{x}(k) = \mathbf{G}^k\mathbf{x}(0)$$

$$= \begin{bmatrix} \frac{1}{3}(1{,}01)^k(10^8) - \frac{0{,}7}{3}(0{,}9494)^k(10^8) \\ \frac{2}{3}(1{,}01)^k(10^8) + \frac{0{,}7}{3}(0{,}9494)^k(10^8) \end{bmatrix}$$

Para $k = 10$, obtemos

$$\mathbf{x}(10) = \begin{bmatrix} \frac{1}{3}(1{,}01)^{10}[1 - 0{,}7(0{,}94)^{10}](10^8) \\ \frac{2}{3}(1{,}01)^{10}[1 + \frac{0{,}7}{2}(0{,}94)^{10}](10^8) \end{bmatrix}$$

$$= \begin{bmatrix} 22{,}94 \times 10^6 \\ 87{,}46 \times 10^6 \end{bmatrix}$$

Portanto, a cidade B terá uma população de 22,94 milhões no ano $k = 10$.

Note que quando k tende a infinito, $\lim_{k\to\infty} (0{,}9494)^k \to 0$. Portanto,

$$\lim_{k\to\infty} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} = \lim_{k\to\infty} \begin{bmatrix} \frac{1}{3}(1{,}01)^k(10^8) \\ \frac{2}{3}(1{,}01)^k(10^8) \end{bmatrix}$$

A razão da população da cidade B para a do resto do país A se torna 1 a 2, ou um terço da população do país A vai morar na cidade B.

PROBLEMAS

Problema B.14.1 Considere o sistema ou circuito visto na Fig. 14.17. Escolhendo v_c e i_L como as variáveis de estado, obtenha a equação de estado do sistema.

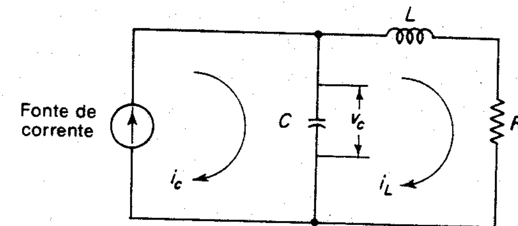


Fig. 14.17 Sistema elétrico.

Problema B.14.2 Considere o sistema descrito por

$$\dddot{y} + 3\ddot{y} + 2\dot{y} = u$$

Determine uma representação de espaço de estados do sistema. Escolha as variáveis de estado de tal forma que a matriz de coeficientes do vetor de estado seja diagonal.

Problema B.14.3 Obtenha uma representação de espaço de estados para o seguinte sistema:

$$\frac{Y(s)}{U(s)} = \frac{2(s+3)}{(s+1)(s+2)}$$

Problema B.14.4 Determine uma representação de espaço de estados do sistema térmico visto na Fig. 14.18. Para simplificar a derivação, suponha o seguinte: A massa de metal do recipiente interno é pequena e sua capacitância térmica é desprezível, assim como a perda de

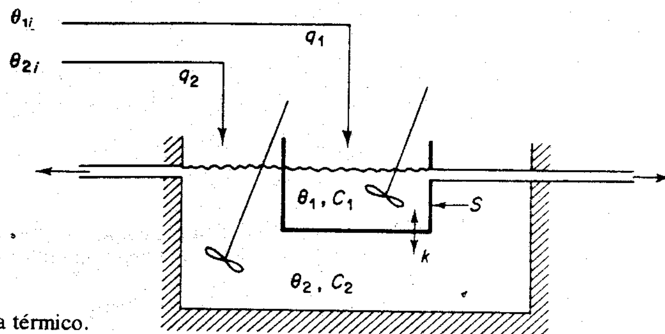


Fig. 14.18 Sistema térmico.

calor do recipiente externo. Além do mais, a perda de calor a partir da superfície livre também é desprezível. Supõe-se que o fluido no recipiente interno é bem misturado a todo instante. O mesmo é verdade para o fluido no recipiente externo. No diagrama

θ_{1i} = temperatura do líquido que entra no recipiente interno, em °F
θ_{2i} = temperatura do líquido que entra no recipiente externo, em °F
θ_1 = temperatura do líquido no recipiente interno, em °F
θ_2 = temperatura do líquido no recipiente externo, em °F
C_1 = capacidade térmica do líquido no recipiente interno, em Btu/°F
C_2 = capacitância térmica do líquido no recipiente externo, em Btu/°F
q_1 = entrada de calor no recipiente interno, em Btu/min°F
q_2 = entrada de calor no recipiente externo, em Btu/min°F
S = área da superfície de transporte de calor, em in²
k = coeficiente de transferência de calor, em Btu/in² min °F

Suponha que θ_1 e θ_2 são as variáveis de estado e que θ_{1i} e θ_{2i} as variáveis de entrada.

Problema B.14.5 Determine $x_1(t)$ e $x_2(t)$ do sistema descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

onde as condições iniciais são

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Problema B.14.6 Dada a equação de um sistema

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Determine a solução em termos das condições iniciais $x_1(0)$, $x_2(0)$ e $x_3(0)$.

Problema B.14.7 Mostre que a solução da seguinte equação diferencial matricial:

$$\frac{d\mathbf{X}}{dt} = \mathbf{AX} + \mathbf{XB}, \qquad \mathbf{X}(0) = \mathbf{C}$$

é dada por

$$\mathbf{X} = e^{\mathbf{A}t}\mathbf{C}e^{\mathbf{B}t}$$

Problema B.14.8 Obtenha a resposta do seguinte sistema:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix} [u], \qquad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

onde

$$\begin{aligned} u(t) &= 0 \quad \text{para } t < 0 \\ &= e^{-t} \quad \text{para } t \geq 0 \end{aligned}$$

Problema B.14.9 Considere o sistema descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -5 & -1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} [u]$$

$$y = [1 \quad 2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Obtenha a função de transferência do sistema.

Problema B.14.10 Considere um sistema de tempo contínuo descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [u]$$

Discretize a equação do sistema e derive uma equação de estado de tempo discreto. Suponha que o período de amostragem é de 2 s.

Problema B.14.11 Obtenha uma representação de espaço de estados do seguinte sistema:

$$\frac{Y(z)}{U(z)} = \frac{z + 2}{z^2 + z + 0{,}16}$$

Problema B.14.12 Considere a equação de diferenças

$$x(k + 2) = x(k + 1) + x(k)$$

onde $x(0) = 0$, $x(1) = 1$. Note que $x(2) = 1$, $x(3) = 2$, $x(4) = 3$, ... A série 0, 1, 1, 2, 3, 5, 8, 13,... é conhecida como a série de Fibonacci.

Obtenha a solução geral $x(k)$. Mostre que o valor limite de $x(k+1)/x(k)$ quando k tende a infinito é $(\sqrt{5} + 1)/2$, ou aproximadamente 1,62.

15

Análise de Estabilidade de Liapunov

15.1 INTRODUÇÃO

Para um dado sistema de controle, a estabilidade é geralmente a coisa mais importante a ser determinada. Se o sistema é linear e invariante no tempo, temos à disposição vários critérios de estabilidade. Entre eles o critério de estabilidade de Nyquist, o critério de estabilidade de Routh etc. Se o sistema é não linear, ou linear mas variando no tempo, então tais critérios de estabilidade não se aplicam. Embora uma técnica que emprega o gráfico de estabilidade de Nyquist possa ser aplicada a um grupo especial de sistemas não lineares, como descrito no Cap. 11, o método da função descritiva para a determinação da estabilidade é apenas aproximado. A análise de estabilidade baseada no método do plano de fase apresentada no Cap. 12 se aplica apenas para sistemas de primeira e segunda-ordem.

O segundo método de Liapunov (que também é chamado método direto de Liapunov) a ser apresentado neste capítulo é o método mais geral para a determinação da estabilidade de sistemas não lineares e/ou sistemas variantes no tempo. O método se aplica para sistemas de qualquer ordem.

Usando o segundo método de Liapunov, podemos determinar a estabilidade de um sistema sem resolver as equações de estado. Isto é uma vantagem porque a solução de equações de estado não lineares e/ou variando no tempo é geralmente muito difícil.

Embora o segundo método de Liapunov requeira uma considerável experiência e engenhosidade, ele pode responder à pergunta sobre a estabilidade de sistemas não lineares quando outros métodos falham.

O objetivo deste capítulo é o de apresentar o segundo método de Liapunov e ilustrar suas aplicações na análise de estabilidade tanto de sistemas lineares quanto não lineares.

Resumo do capítulo. Na Seção 15.2 apresentamos tópicos preliminares como definições dos vários tipos de estabilidade de sistemas e os conceitos de funções escalares definidas. A Seção 15.3 apresenta o teorema principal de estabilidade do segundo método e introduz a função de Liapunov. A Seção 15.4 discute a aplicação do segundo método de Liapunov para a análise de estabilidade de sistemas lineares invariantes no tempo. Em seguida, na Seção 15.5 apresentamos uma técnica útil baseada no segundo método para estimar o comportamento transitório de sistemas. A Seção 15.6 discute a análise de estabilidade de sistemas não lineares. Apresentamos dois métodos para construir funções de Liapunov para tais sistemas. Finalmente, a Seção 15.7 resume comentários conclusivos para o capítulo.

15.2 DEFINIÇÕES

Nesta seção, inicialmente apresentamos definições de um sistema, de um estado de equilíbrio, estabilidade, estabilidade assintótica, e instabilidade. Em seguida definimos a característica de definição, semidefinição e indefinição de funções escalares.

Sistema. O sistema considerado neste capítulo é definido por

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \qquad (15\text{-}1)$$

onde $\mathbf{x}$ é um vetor de estado (vetor n-dimensional) e $\mathbf{f}(\mathbf{x}, t)$ é um vetor n-dimensional cujos elementos são funções de $x_1, x_2, \ldots, x_n$, e t. Supomos que o sistema da Eq. (15.1) tem uma única solução começando em uma dada condição inicial.

Denotaremos a solução da Eq. (15.1) como $\boldsymbol{\phi}(t; \mathbf{x}_0, t_0)$, onde $\mathbf{x} = \mathbf{x}_0$ em $t = t_0$ e t é o tempo observado. Portanto,

$$\boldsymbol{\phi}(t_0; \mathbf{x}_0, t_0) = \mathbf{x}_0$$

Estado de equilíbrio. No sistema da Eq. (15.1), um estado $\mathbf{x}_e$, onde

$$\mathbf{f}(\mathbf{x}_e, t) = \mathbf{0} \qquad \text{para todo } t \qquad (15\text{-}2)$$

é chamado de um estado de equilíbrio do sistema. Se o sistema é linear invariante no tempo, ou seja, se $\mathbf{f}(\mathbf{x}, t) = \mathbf{A}\mathbf{x}$, então há apenas um estado de equilíbrio se $\mathbf{A}$ é não singular e um número infinito de estados de equilíbrio se $\mathbf{A}$ é singular. Para sistemas não lineares, pode haver um ou mais estados de equilíbrio. Estes estados correspondem às soluções constantes do sistema ($\mathbf{x} = \mathbf{x}_e$ para todo t). A determinação dos estados de equilíbrio não envolve a solução das equações diferenciais do sistema, Eq. (15.1), mas apenas a solução da Eq. (15.2).

Qualquer estado de equilíbrio isolado (isto é, isolado um do outro) pode ser deslocado para a origem das coordenadas, ou $\mathbf{f}(\mathbf{0}, t) = \mathbf{0}$, através de uma translação de coordenadas. Neste capítulo, trataremos da análise de estabilidade apenas de tais estados.

Estabilidade no sentido de Liapunov. A seguir, denotaremos uma região esférica de raio k ao redor de um estado de equilíbrio $\mathbf{x}_e$ como

$$\|\mathbf{x} - \mathbf{x}_e\| \leq k$$

onde $\|\mathbf{x} - \mathbf{x}_e\|$ é chamada de norma Euclidiana e definida por

$$\|\mathbf{x} - \mathbf{x}_e\| = [(x_1 - x_{1e})^2 + (x_2 - x_{2e})^2 + \cdots + (x_n - x_{ne})^2]^{1/2}$$

Façamos $S(\delta)$ consistir em todos os pontos tais que

$$\|\mathbf{x}_0 - \mathbf{x}_e\| \leq \delta$$

e seja $S(\epsilon)$ o conjunto de todos os pontos tais que

$$\|\boldsymbol{\phi}(t; \mathbf{x}_0, t_0) - \mathbf{x}_e\| \leq \epsilon \qquad \text{para todo } t \geqslant t_0$$

Um estado de equilíbrio $\mathbf{x}_e$ do sistema da Eq. (15.1) é dito estável no sentido de Liapunov se, correspondendo a cada $S(\epsilon)$, há um $S(\delta)$ tal que trajetórias partindo de $S(\delta)$ não saem de $S(\epsilon)$ quando t aumenta indefinidamente. O número real δ depende de ϵ e, em geral, também de t_0. Se δ não depende de t_0, o estado de equilíbrio é dito uniformemente estável.

O que foi especificado aqui é que primeiro se escolhe a região $S(\epsilon)$, e para cada $S(\epsilon)$ deve haver uma região $S(\delta)$ tal que trajetórias partindo dentro de $S(\delta)$ não abandonam $S(\epsilon)$ quando t aumenta indefinidamente.

Estabilidade assintótica. Um estado de equilíbrio $\mathbf{x}_e$ do sistema da Eq. (15.1) é dito assintoticamente estável se ele é estável no sentido de Liapunov e se toda solução partindo de dentro de $S(\delta)$ converge, sem sair de $S(\epsilon)$, para $\mathbf{x}_e$ quando t aumenta indefinidamente.

Na prática, a estabilidade assintótica é mais importante do que a mera estabilidade. Também, como estabilidade assintótica é um conceito local, a simples confirmação de estabilidade assintótica pode não significar que o sistema vai operar adequadamente. Algum conhecimento do tamanho da máxima região de estabilidade assintótica é geralmente necessário. Esta região é chamada o domínio de atração. Este domínio é a parte do espaço de estados em que se originam trajetórias assintoticamente estáveis. Em outras palavras, toda trajetória que se origina no domínio de atração é assintoticamente estável.

Estabilidade assintótica global. Se a estabilidade assintótica vale para todos os estados (todos os pontos do espaço de estados) a partir dos quais se originam trajetórias, o estado de equilíbrio é dito assintoticamente estável globalmente. Isto é, o estado de equilíbrio $\mathbf{x}_e$ do sistema dado pela Eq. (15.1) é dito assintoticamente estável globalmente se ele é estável e se toda solução converge para $\mathbf{x}_e$ quando t aumenta indefinidamente. Obviamente uma condição necessária para estabilidade assintótica global é que haja apenas um estado de equilíbrio em todo o espaço de estados.

Em problemas de engenharia de controle, uma característica desejável é a estabilidade assintótica global. Se o estado de equilíbrio não é assintoticamente estável globalmente, então o problema se torna o de determinar a maior região de estabilidade assintótica. Isto é geralmente muito difícil. Entretanto, para todos os fins práticos, é suficiente determinar uma região de estabilidade assintótica suficientemente grande tal que nenhuma perturbação a exceda.

Instabilidade. Um estado de equilíbrio $\mathbf{x}_e$ é dito instável se para algum número real $\epsilon > 0$ e qualquer número real $\delta > 0$, tão pequenos quanto se queira, sempre há um estado $\mathbf{x}_0$ em $S(\delta)$ tal que a trajetória partindo deste estado abandona $S(\epsilon)$.

Representação gráfica de estabilidade, estabilidade assintótica e instabilidade. Uma representação gráfica destas definições vistas esclarecerá seus significados.

Vamos considerar o caso bidimensional. As Figs. 15.1 (a), (b) e (c) mostram estados de equilíbrio e trajetórias típicas correspondendo à estabilidade, estabilidade assintótica e instabilidade, respectivamente. Na Fig. 15.1 (a), (b) ou (c) a região $S(\delta)$ contém o estado inicial $\mathbf{x}_0$ e a região $S(\epsilon)$ corresponde à região contendo a trajetória começando em $\mathbf{x}_0$.

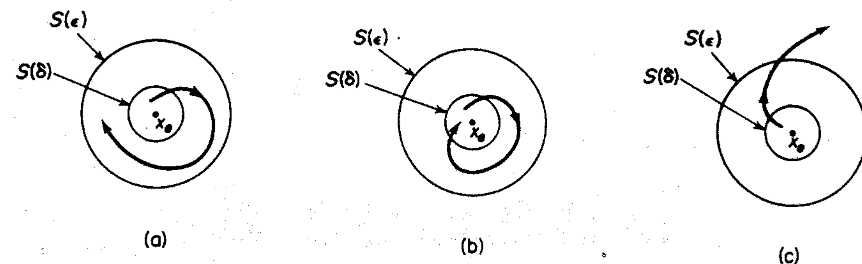


Fig. 15.1 (a) Estado de equilíbrio estável e uma trajetória representativa; (b) estado de equilíbrio assintoticamente estável e uma trajetória representativa; (c) estado de equilíbrio instável e uma trajetória representativa.

Note que as definições dadas não especificam a região exata de condições iniciais permitidas. Portanto a definição se aplica à vizinhança do estado de equilíbrio, a não ser que $S(\epsilon)$ corresponda à totalidade do plano de estados.

Note que na Fig. 15.1(c), a trajetória sai de $S(\epsilon)$ e implica que o estado de equilíbrio é instável. Não podemos entretanto dizer que a trajetória irá para o infinito, pois ela pode tender a um ciclo limite fora da região $S(\epsilon)$. (Se um sistema linear invariante no tempo é instável, as trajetórias começando perto de um estado de equilíbrio instável se dirigem ao infinito. Mas no caso de sistemas não lineares, isto não é necessariamente verdade.)

O conhecimento destas definições é o requisito mínimo para o entendimento da análise de estabilidade de sistemas lineares e não lineares conforme apresentado neste capítulo. Note que estas definições não são as únicas que definem conceitos de estabilidade de um estado de equilíbrio. De fato, várias outras formas são encontradas na literatura. Por exemplo, na teoria de controle clássica, apenas os sistemas assintoticamente estáveis são chamados de sistemas estáveis, e aqueles que são estáveis no sentido de Liapunov, mas não assintoticamente estáveis, são chamados de instáveis.

Funções escalares positivas definidas. Uma função escalar $V(\mathbf{x})$ é dita *positiva definida* em uma região $\boldsymbol{\Omega}$ (que inclui a origem do espaço de estados) se $V(\mathbf{x}) > 0$ para todos os estados não nulos $\mathbf{x}$ na região $\boldsymbol{\Omega}$ e $V(\mathbf{0}) = 0$.

Uma função variando no tempo $V(\mathbf{x}, t)$ é dita positiva definida em uma região $\boldsymbol{\Omega}$ (que inclui a origem do espaço de estados) se é limitada inferiormente por uma função positiva definida $V(\mathbf{x})$ tal que

$$V(\mathbf{x}, t) > V(\mathbf{x}) \qquad \text{para todo } t \geqslant t_0$$
$$V(\mathbf{0}, t) = 0 \qquad \text{para todo } t \geqslant t_0$$

Funções escalares negativas definidas. Uma função escalar $V(\mathbf{x})$ é dita *negativa definida* se $-V(\mathbf{x})$ é positiva definida.

Funções escalares positivas semidefinidas. Uma função escalar $V(\mathbf{x})$ é dita *positiva semidefinida* se é positiva para todos os estados na região $\boldsymbol{\Omega}$, exceto na origem, e em certos outros estados, onde ela é nula.

Funções escalares negativas semidefinidas. Uma função escalar $V(\mathbf{x})$ é dita *negativa semidefinida* se $-V(\mathbf{x})$ é positiva semidefinida.

Funções escalares indefinidas. Uma função escalar é dita *indefinida* se na região Ω assume tanto valores positivos como negativos, independentemente de quão pequena é a região Ω.

Exemplo 15.1 Neste exemplo, damos várias funções escalares e suas classificações de acordo com as definições vistas. Aqui supomos que $\mathbf{x}$ é um vetor bidimensional.

1.	$V(\mathbf{x}) = x_1^2 + 2x_2^2$	positiva definida
2.	$V(\mathbf{x}) = (x_1 + x_2)^2$	positiva semidefinida
3.	$V(\mathbf{x}) = -x_1^2 - (3x_1 + 2x_2)^2$	negativa definida
4.	$V(\mathbf{x}) = x_1 x_2 + x_2^2$	indefinida
5.	$V(\mathbf{x}) = x_1^2 + \dfrac{2x_2^2}{1 + x_2^2}$	positiva definida

Formas quadráticas. Uma classe de funções escalares que têm um papel importante na análise de estabilidade baseada no segundo método de Liapunov é a forma quadrática. Um exemplo é

$$V(\mathbf{x}) = \mathbf{x}'\mathbf{P}\mathbf{x}$$

$$= [x_1 \quad x_2 \quad \cdots \quad x_n]\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{12} & p_{22} & \cdots & p_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ p_{1n} & p_{2n} & \cdots & p_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

Note que $\mathbf{P}$ é real e simétrica.

Para determinar se a forma quadrática $V(\mathbf{x})$ é positiva definida pode-se utilizar o critério de Sylvester, o qual diz que as condições necessárias e suficientes para que a forma quadrática $V(\mathbf{x})$ seja positiva definida são que todos os menores principais* sucessivos de $\mathbf{P}$ sejam positivos: isto é,

$$p_{11} > 0, \qquad \begin{vmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{vmatrix} > 0, \ldots,$$

$$\begin{vmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{12} & p_{22} & \cdots & p_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ p_{1n} & p_{2n} & \cdots & p_{nn} \end{vmatrix} > 0$$

---

*N. do T.: Nem todos os menores principais, mas apenas aqueles com a forma indicada no texto. Por exemplo:

$\begin{bmatrix} p_{11} & p_{13} \\ p_{31} & p_{33} \end{bmatrix}$ *não* nos interessa.

$V(\mathbf{x}) = \mathbf{x}'\mathbf{P}\mathbf{x}$ é positiva semidefinida se $\mathbf{P}$ é singular e todos os menores principais são não negativos.

$V(\mathbf{x})$ é negativa definida se $-V(\mathbf{x})$ é positiva definida. De forma semelhante, $V(\mathbf{x})$ é negativa semidefinida se $-V(\mathbf{x})$ é positiva semidefinida.

Exemplo 15.2 Mostre que a seguinte forma quadrática é positiva definida:

$$V(\mathbf{x}) = 10x_1^2 + 4x_2^2 + x_3^2 + 2x_1x_2 - 2x_2x_3 - 4x_1x_3$$

A forma quadrática $V(\mathbf{x})$ pode ser escrita como

$$V(\mathbf{x}) = \mathbf{x}'\mathbf{P}\mathbf{x} = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 10 & 1 & -2 \\ 1 & 4 & -1 \\ -2 & -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Aplicando o critério de Sylvester, obtemos

$$10 > 0, \qquad \begin{vmatrix} 10 & 1 \\ 1 & 4 \end{vmatrix} > 0, \qquad \begin{vmatrix} 10 & 1 & -2 \\ 1 & 4 & -1 \\ -2 & -1 & 1 \end{vmatrix} > 0$$

Como todos os menores principais sucessivos da matriz $\mathbf{P}$ são positivos, $V(\mathbf{x})$ é positiva definida.

15.3 O SEGUNDO MÉTODO DE LIAPUNOV

Em 1892, A. M. Liapunov apresentou dois métodos (chamados de primeiro e segundo métodos) para determinar a estabilidade de sistemas dinâmicos descritos por equações diferenciais ordinárias.

O primeiro método consiste em todos procedimentos em que a forma explícita das soluções das equações diferenciais é usada para a análise.

O segundo método, por outro lado, não requer as soluções das equações diferenciais. Portanto, é bastante conveniente para a análise de estabilidade de sistemas não lineares, para os quais soluções exatas podem não ser determináveis.

Segundo método de Liapunov. Da teoria clássica da mecânica, sabemos que um sistema vibratório é estável se sua energia total (uma função positiva definida) é continuamente decrescente (o que significa que a derivada temporal da energia total deve ser negativa definida) até que um estado de equilíbrio é alcançado.

O segundo método de Liapunov é baseado em uma generalização deste fato: se o sistema tem um estado de equilíbrio assintoticamente estável, então a energia armazenada do sistema deslocado dentro do domínio de atração decresce com o passar do tempo, até que finalmente assume um valor mínimo no estado de equilíbrio. Para sistemas puramente matemáticos, entretanto, não há uma maneira simples de definir uma "função de energia". Para contornar esta dificuldade, Liapunov introduziu a chamada função de Liapunov, uma função de energia fictícia. A idéia é, entretanto, mais geral que a de energia, e é aplicável de forma mais ampla. De fato, qualquer função escalar que satisfaz as hipóteses dos teoremas de estabilidade de Liapunov (veja Teoremas 15.1 e 15.2) pode servir como funções de Liapunov. (Para sistemas simples, podemos ser capazes de adivinhar

funções de Liapunov adequadas; mas, para um sistema complicado, achar uma função de Liapunov pode ser bastante difícil.)

As funções de Liapunov dependem de $x_1, x_2, \ldots x_n$, e de t. Denotamos estas por $V(x_1, x_2, \ldots x_n, t)$ ou simplesmente por $V(\mathbf{x}, t)$. Se as funções de Liapunov não incluírem o tempo t explicitamente, então denotamos por $V(x_1, x_2, \ldots x_n)$, ou $V(\mathbf{x})$. No segundo método de Liapunov, o comportamento do sinal de $V(\mathbf{x}, t)$ e o de sua derivada temporal $\dot{V}(\mathbf{x}, t) = dV(\mathbf{x}, t)/dt$ nos fornecem informação sobre a estabilidade, estabilidade assintótica, ou instabilidade de um estado de equilíbrio sem que haja a necessidade de se resolver diretamente as equações. (Isto se aplica tanto para sistemas lineares quanto não lineares.)

Teorema principal sobre estabilidade de Liapunov. Pode-se provar que, se uma função escalar $V(\mathbf{x})$, onde $\mathbf{x}$ é um vetor n-dimensional, é positiva definida, então os estados que satisfazem

$$V(\mathbf{x}) = C$$

onde C é uma constante positiva, se localizam em uma hiper-superfície fechada no espaço de estados n-dimensional, pelo menos nas vizinhanças da origem. Se $V(\mathbf{x}) \to \infty$ quando $\|\mathbf{x}\| \to \infty$, então tais superfícies fechadas se estendem sobre todo o espaço de estados. A hiper-superfície $V(\mathbf{x}) = C_1$ está inteiramente dentro da hiper-superfície $V(\mathbf{x}) = C_2$ se $C_1 < C_2$.

Para um dado sistema, caso se possa achar uma função escalar positiva definida tal que sua derivada temporal ao longo de uma trajetória é sempre negativa, então, quando o tempo aumenta, $V(\mathbf{x})$ assume valores cada vez mais baixos de C. Quando o tempo aumenta, $V(\mathbf{x})$ finalmente vai a zero, e portanto $\mathbf{x}$ também vai a zero. Isso implica que a origem do espaço de estados é assintoticamente estável. O teorema principal de Liapunov sobre estabilidade, que é uma generalização dos fatos agora descritos, fornece uma condição suficiente para estabilidade assintótica. Este teorema pode ser enunciado como segue:

Teorema 15.1 Suponha que um sistema é descrito por

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$$

onde

$$\mathbf{f}(\mathbf{0}, t) = \mathbf{0} \qquad \text{para todo } t$$

Se há uma função escalar $V(\mathbf{x}, t)$ tendo primeiras derivadas parciais contínuas e satisfazendo as seguintes condições:

1. $V(\mathbf{x}, t)$ é positiva definida
2. $\dot{V}(\mathbf{x}, t)$ é negativa definida

então o estado de equilíbrio na origem é assintoticamente estável uniformemente.

Se, além disso, $V(\mathbf{x}, t) \to \infty$ para $\|\mathbf{x}\| \to \infty$, então o estado de equilíbrio na origem é assintoticamente estável uniformemente e de de forma global.

Não forneceremos detalhes da prova do teorema. (A prova segue diretamente da definição de estabilidade assintótica. Para detalhes, veja, por exemplo, o Cap. 8 da Referência O-1.)

Exemplo 15.3 Considere o sistema descrito por

$$\dot{x}_1 = x_2 - x_1(x_1^2 + x_2^2)$$
$$\dot{x}_2 = -x_1 - x_2(x_1^2 + x_2^2)$$

É claro que a origem ($x_1 = 0, x_2 = 0$) é o único estado de equilíbrio. Determine a sua estabilidade.

Se definirmos uma função escalar $V(\mathbf{x})$ como

$$V(\mathbf{x}) = x_1^2 + x_2^2$$

que é positiva definida, então a derivada temporal de $V(\mathbf{x})$ ao longo de qualquer trajetória é

$$\dot{V}(\mathbf{x}) = 2x_1\dot{x}_1 + 2x_2\dot{x}_2$$
$$= -2(x_1^2 + x_2^2)^2$$

que é negativa definida. Isto mostra que $V(\mathbf{x})$ é continuamente decrescente ao longo de qualquer trajetória; portanto $V(\mathbf{x})$ é uma função de Liapunov. E como $V(\mathbf{x})$ se torna infinito para um desvio infinito a partir do estado de equilíbrio, pelo Teorema 15.1, o estado de equilíbrio na origem do sistema é assintoticamente estável globalmente.

Note que se fizermos $V(\mathbf{x})$ assumir valores constantes, $0, C_1, C_2, \ldots$ $(0 < C_1 < C_2 \ldots)$, então $V(\mathbf{x}) = 0$ corresponde à origem do plano de estados e $V(\mathbf{x}) = C_1$, $V(\mathbf{x}) = C_2, \ldots$ descrevem círculos que não se interceptam, contendo a origem do plano de estados, como visto na Fig. 15.2. Note também que como $V(\mathbf{x})$ não é limitada radialmente, ou $V(\mathbf{x}) \to \infty$ quando $\|\mathbf{x}\| \to \infty$, os círculos se estendem sobre todo o plano de estados.

Como o círculo $V(\mathbf{x}) = C_k$ fica completamente dentro do círculo $V(\mathbf{x}) = C_{k+1}$, uma trajetória representativa cruza o limite dos contornos de V de fora para dentro. Daí, a interpretação geométrica de uma função de Liapunov pode ser enunciada como segue: $V(\mathbf{x})$ é uma medida da distância do estado $\mathbf{x}$ a partir da origem do espaço de estados. Se a distância entre a origem e o estado instantâneo $\mathbf{x}(t)$ é continuamente decrescente quando t aumenta (isto é, $\dot{V}(\mathbf{x}(t)) < 0$), então $\mathbf{x}(t) \to \mathbf{0}$. Como uma função de Liapunov $V(\mathbf{x})$ é interpretada como definindo uma distância a partir da origem do espaço de estados, sua derivada pode ser utilizada para fornecer uma estimativa quantitativa da velocidade com que a origem está sendo alcançada. (Para detalhes, veja Seção 15.5.)

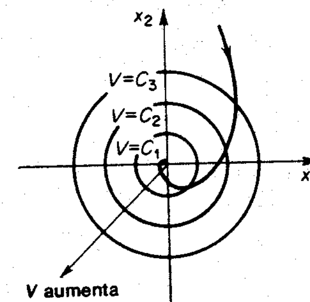


Fig. 15.2 Curvas de contorno de V constante e uma trajetória representativa.

Comentário. Embora o Teorema 15.1 seja um teorema básico do segundo método, ele é um tanto restritivo porque $\dot{V}(\mathbf{x}, t)$ deve ser negativa definida. Se, entretanto, uma restrição adicional é imposta sobre $\dot{V}(\mathbf{x}, t)$, que ela não se anule ao longo de qualquer trajetória exceto na origem, então é possível substituir a condição de $\dot{V}(\mathbf{x}, t)$ negativa definida pela condição de $\dot{V}(\mathbf{x}, t)$ negativa semidefinida.

Teorema 15.2 Suponha que um sistema é descrito por

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$$

onde

$$\mathbf{f}(\mathbf{0}, t) = \mathbf{0} \quad \text{para todo } t \geq t_0$$

Se há uma função escalar $V(\mathbf{x}, t)$ tendo primeiras derivadas parciais contínuas e satisfazendo as seguintes condições:

1. $V(\mathbf{x}, t)$ é positiva definida
2. $\dot{V}(\mathbf{x}, t)$ é negativa semidefinida
3. $\dot{V}(\phi(t; \mathbf{x}_0, t_0), t)$ não se anulando em $t \geqslant t_0$ para qualquer t_0 e qualquer $\mathbf{x}_0 \neq \mathbf{0}$, onde $\phi(t; \mathbf{x}_0, t_0)$ denota a trajetória ou solução partindo de $\mathbf{x}_0$ em t_0.

então o estado de equilíbrio na origem do sistema é uniforme e assintoticamente estável globalmente.

Note que se $\dot{V}(\mathbf{x}, t)$ não é negativa definida, mas apenas negativa semidefinida, então a trajetória de um ponto representativo pode-se tornar tangente a alguma superfície particular $V(\mathbf{x}, t) = C$. Entretanto, como $\dot{V}(\phi(t; \mathbf{x}_0, t_0), t)$ não se anula em $t \geqslant t_0$ para qualquer t_0 e qualquer $\mathbf{x}_0 \neq \mathbf{0}$, o ponto representativo não pode permanecer no ponto tangente (o ponto que corresponde a $\dot{V}(\mathbf{x}, t) = 0$) e portanto deve-se mover em direção à origem.

Se, entretanto, há uma função escalar positiva definida $V(\mathbf{x}, t)$ tal que $\dot{V}(\mathbf{x}, t)$ é identicamente zero, então o sistema pode permanecer em um ciclo limite. O estado de equilíbrio na origem, neste caso, é dito estável no sentido de Liapunov.

Exemplo 15.4 Considere o seguinte sistema:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

É claro que o único estado de equilíbrio é a origem, $\mathbf{x} = \mathbf{0}$. Determine a estabilidade deste estado.

Vamos escolher a seguinte função escalar como uma possível função de Liapunov:

$$V(\mathbf{x}) = 2x_1^2 + x_2^2 = \text{positiva definida}$$

Então $\dot{V}(\mathbf{x})$ se torna

$$\dot{V}(\mathbf{x}) = 4x_1\dot{x}_1 + 2x_2\dot{x}_2 = 2x_1x_2 - 2x_2^2$$

$\dot{V}(\mathbf{x})$ é indefinida. Isto implica que este particular $V(\mathbf{x})$ não é uma função de Liapunov, e portanto a estabilidade não pode ser determinada através do seu uso. (Como os autovalores da matriz dos coeficientes são $(-1 + j\sqrt{3})/2$ e $(-1 - j\sqrt{3})/2$, é claro que a origem do sistema é estável. Isto significa que não escolhemos uma função de Liapunov adequada.)

Se escolhermos a seguinte função escalar como uma possível função de Liapunov.

$$V(\mathbf{x}) = x_1^2 + x_2^2 = \text{positiva definida}$$

então

$$\dot{V}(\mathbf{x}) = 2x_1\dot{x}_1 + 2x_2\dot{x}_2 = -2x_2^2$$

que é negativa semidefinida. Se $\dot{V}(\mathbf{x})$ deve se anular para todo $t \geqslant t_1$, então x_2 deve ser zero para todo $t \geqslant t_1$. Isto requer que $\dot{x}_2 = 0$ para $t \geqslant t_1$. Como

$$\dot{x}_2 = -x_1 - x_2$$

x_1 deve também ser igual a zero para $t \geqslant t_1$. Isto significa que $\dot{V}(x)$ se anula apenas na origem. Portanto, pelo Teorema 15.2, o estado de equilíbrio na origem é assintoticamente estável globalmente.

Para mostrar que uma escolha diferente de uma função de Liapunov fornece a mesma informação de estabilidade, vamos escolher a seguinte função escalar como outra possível função de Liapunov:

$$V(\mathbf{x}) = \tfrac{1}{2}[(x_1 + x_2)^2 + 2x_1^2 + x_2^2] = \text{positiva definida}$$

então $\dot{V}(\mathbf{x})$ se torna

$$\dot{V}(\mathbf{x}) = (x_1 + x_2)(\dot{x}_1 + \dot{x}_2) + 2x_1\dot{x}_1 + x_2\dot{x}_2 = -(x_1^2 + x_2^2)$$

que é negativa definida. Como $V(\mathbf{x}) \to \infty$ quando $\|\mathbf{x}\| \to \infty$, pelo Teorema 15.1, o estado de equilíbrio na origem é assintoticamente estável globalmente.

Note que como os teoremas de estabilidade do segundo método requerem $V(\mathbf{x})$ positiva definida, freqüentemente (mas não sempre) escolhemos $V(\mathbf{x})$ como uma forma quadrática em $\mathbf{x}$. (Note que a função positiva definida mais simples é uma forma quadrática.) Em seguida examinamos se $\dot{V}(\mathbf{x})$ é pelo menos negativa semidefinida.

Instabilidade. Se um estado de equilíbrio $\mathbf{x} = \mathbf{0}$ de um sistema é instável, então existe uma função escalar $W(\mathbf{x}, t)$ que determina a instabilidade do estado de equilíbrio. Apresentaremos um teorema sobre instabilidade sem uma demonstração.

Teorema 15.3 Suponha que um sistema é descrito por

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$$

onde

$$\mathbf{f}(\mathbf{0}, t) = \mathbf{0} \quad \text{para todo } t \geqslant t_0$$

Se há uma função escalar $W(\mathbf{x}, t)$, tendo primeiras derivadas parciais contínuas e satisfazendo as seguintes condições:

1. $W(\mathbf{x}, t)$ é positiva definida em uma certa região ao redor da origem
2. $\dot{W}(\mathbf{x}, t)$ é positiva definida na mesma região

então o estado de equilíbrio na origem é instável.

15.4 ANÁLISE DE ESTABILIDADE DE SISTEMAS LINEARES

Há muitas abordagens para a investigação da estabilidade assintótica de sistemas lineares inverianTes no tempo. Por exemplo, para um sistema de tempo contínuo

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$$

a condição necessária e suficiente para a estabilidade assintótica da origem do sistema pode ser enunciada como todos os autovalores de $\mathbf{A}$ tendo partes reais

negativas, ou os zeros do polinômio característico

$$|s\mathbf{I} - \mathbf{A}| = s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n$$

tendo partes reais negativas. De modo semelhante, para um sistema de tempo discreto

$$\mathbf{x}(k+1) = \mathbf{G}\mathbf{x}(k)$$

a condição necessária e suficiente para a estabilidade assintótica da origem pode ser enunciada como todos os autovalores de **G** tendo módulo menor que a unidade, ou que os zeros do polinômio característico

$$|z\mathbf{I} - \mathbf{G}| = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n$$

estão dentro do círculo unitário centrado na origem do plano z.

A determinação dos autovalores se torna difícil ou impossível no caso de sistemas de ordem alta ou se alguns dos coeficientes do polinômio característico são não-numéricos. Neste caso, o critério de estabilidade de Routh pode ser aplicado convenientemente. Uma alternativa para este método é baseada no segundo método de Liapunov. A abordagem de Liapunov é algébrica e não requer a fatoração do polinômio característico. Além do mais este método pode ser usado para estimar o comportamento de resposta transitória do sistema (veja Seção 15.5) e pode ser aplicado no projeto do sistema (veja Seção 16.6). A finalidade desta seção é apresentar a abordagem de Liapunov na análise de estabilidade de sistemas lineares invariantes no tempo.

Análise de estabilidade de Liapunov para sistemas lineares invariantes no tempo. Considere o seguinte sistema linear invariante no tempo:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{15-3}$$

onde **x** é um vetor de estado (vetor n-dimensional) e **A** é uma matriz constante $n \times n$. Supomos que **A** é não singular. Então o único estado de equilíbrio é a origem $\mathbf{x} = \mathbf{0}$. A estabilidade do estado de equilíbrio do sistema linear invariante no tempo pode ser investigada facilmente usando-se o segundo método de Liapunov.

Para o sistema definido pela Eq. (15.3), vamos escolher uma possível função de Liapunov como

$$V(\mathbf{x}) = \mathbf{x}^*\mathbf{P}\mathbf{x}$$

onde **P** é uma matriz positiva definida Hermitiana. (Se **x** é um vetor real, então **P** pode ser escolhida como uma matriz real simétrica positiva definida.) A derivada de $V(\mathbf{x})$ em relação ao tempo ao longo de qualquer trajetória é

$$\begin{aligned}\dot{V}(\mathbf{x}) &= \dot{\mathbf{x}}^*\mathbf{P}\mathbf{x} + \mathbf{x}^*\mathbf{P}\dot{\mathbf{x}} \\ &= (\mathbf{A}\mathbf{x})^*\mathbf{P}\mathbf{x} + \mathbf{x}^*\mathbf{P}\mathbf{A}\mathbf{x} \\ &= \mathbf{x}^*\mathbf{A}^*\mathbf{P}\mathbf{x} + \mathbf{x}^*\mathbf{P}\mathbf{A}\mathbf{x} \\ &= \mathbf{x}^*(\mathbf{A}^*\mathbf{P} + \mathbf{P}\mathbf{A})\mathbf{x}\end{aligned}$$

Como $V(\mathbf{x})$ foi escolhido positivo definido, requeremos, para estabilidade assintótica, que $\dot{V}(\mathbf{x})$ seja negativa definida. Portanto, requeremos que

$$\dot{V} = -\mathbf{x}^*\mathbf{Q}\mathbf{x}$$

onde

$$\mathbf{Q} = -(\mathbf{A}^*\mathbf{P} + \mathbf{P}\mathbf{A}) = \text{positiva definida}$$

Portanto, para a estabilidade assintótica do sistema da Eq. (15.3), é suficiente que **Q** seja positiva definida. Para um teste para determinar se uma matriz $n \times n$ é positiva definida, aplicamos a condição de Sylvester, que diz que uma condição necessária e suficiente para que a matriz seja positiva definida é que os determinantes de todos os menores principais sucessivos da matriz sejam positivos.

Ao invés de especificar uma matriz positiva definida **P** e examinar se **Q** é ou não positiva definida, é conveniente especificar inicialmente uma matriz positiva definida **Q** e então examinar se **P** determinada de

$$\mathbf{A}^*\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}$$

é positiva definida. Note que **P**, sendo positiva definida, é uma condição necessária e suficiente. Resumiremos o que acabamos de ver na forma de um teorema.

Teorema 15.4 Considere o sistema descrito por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$$

onde **x** é um vetor de estado (vetor n-dimensional) e **A** uma matriz constante $n \times n$ não singular. Uma condição necessária e suficiente para que o estado de equilíbrio $\mathbf{x} = \mathbf{0}$ seja assintoticamente estável globalmente é que, dada uma matriz positiva definida Hermitiana (ou simétrica real) **Q**, exista uma matriz positiva definida Hermitiana (ou real simétrica) **P** tal que

$$\mathbf{A}^*\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}$$

A função escalar $\mathbf{x}^*\mathbf{P}\mathbf{x}$ é uma função de Liapunov para este sistema. (Note que no sistema linear em consideração, se o estado de equilíbrio (a origem) é assintoticamente estável, então ele é assintoticamente estável globalmente.)

Ao aplicar este teorema, várias observações importantes são necessárias.

1. Se $\dot{V}(\mathbf{x}) = -\mathbf{x}^*\mathbf{Q}\mathbf{x}$ não se anula ao longo de qualquer trajetória, então **Q** pode ser escolhida positiva semidefinida.
2. Se escolhermos uma matriz arbitrária **Q** positiva definida [ou uma matriz arbitrária **Q** positiva semidefinida se $\dot{V}(\mathbf{x})$ não se anula ao longo de qualquer trajetória] e resolvermos a equação matricial

$$\mathbf{A}^*\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}$$

para determinar **P**, então **P** positiva definida é uma condição necessária e suficiente para a estabilidade assintótica do estado de equilíbrio $\mathbf{x} = \mathbf{0}$.
3. O resultado final não depende da matriz **Q** escolhida em particular, contanto que ela seja positiva definida (ou positiva semidefinida, conforme o caso).
4. Para determinar os elementos da matriz **P**, equacionamos as matrizes $\mathbf{A}^*\mathbf{P} + \mathbf{P}\mathbf{A}$ e $-\mathbf{Q}$ elemento por elemento. Isto resulta em $n(n+1)/2$ equações

lineares para a determinação dos elementos $p_{ij} = \bar{p}_{ji}$ de **P**. Se chamarmos os autovalores de **A** de $\lambda_1, \lambda_2, \ldots, \lambda_n$, cada um repetido tantas vezes quanto for sua multiplicidade como uma raiz da equação característica, e se para cada soma de duas raízes

$$\lambda_j + \lambda_k \neq 0$$

então os elementos de **P** são univocamente determinados. Note que se a matriz **A** representa um sistema estável, então as somas $\lambda_j + \lambda_k$ são sempre não nulas.

5. Ao se determinar se existe ou não uma matriz Hermitiana positiva definida ou uma matriz **P** real e simétrica, é conveniente escolher **Q = I**, onde **I** é a matriz identidade. Então os elementos de **P** são determinados de

$$\mathbf{A^*P} + \mathbf{PA} = -\mathbf{I}$$

e a matriz **P** é testada para ver se é positiva definida.

Exemplo 15.5 Considere o sistema de segunda-ordem descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

É claro que o estado de equilíbrio é a origem. Determine a estabilidade deste estado.

Vamos supor uma função de Liapunov como tentativa

$$V(\mathbf{x}) = \mathbf{x}'\mathbf{P}\mathbf{x}$$

onde **P** será determinado de

$$\mathbf{A}'\mathbf{P} + \mathbf{PA} = -\mathbf{I}$$

ou

$$\begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Expandindo esta equação matricial, obtemos três equações simultâneas como segue:

$$\begin{aligned} -2p_{12} &= -1 \\ p_{11} - p_{12} - p_{22} &= 0 \\ 2p_{12} - 2p_{22} &= -1 \end{aligned}$$

Resolvendo para determinar p_{11}, p_{12}, p_{13}, obtemos

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$$

Para testar se **P** é positiva definida, conferimos os determinantes dos menores principais sucessivos:

$$p_{11} = \frac{3}{2} > 0, \qquad \begin{vmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{vmatrix} = \begin{vmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{vmatrix} > 0$$

Claramente, **P** é positiva definida. Portanto, o estado de equilíbrio na origem é assintoticamente estável globalmente, e uma função de Liapunov é

$$V = \mathbf{x}'\mathbf{P}\mathbf{x} = \tfrac{1}{2}(3x_1^2 + 2x_1x_2 + 2x_2^2)$$

e

$$\dot{V} = -(x_1^2 + x_2^2)$$

Exemplo 15.6 Determine a gama de estabilidade para o ganho K do sistema visto na Fig. 15.3.

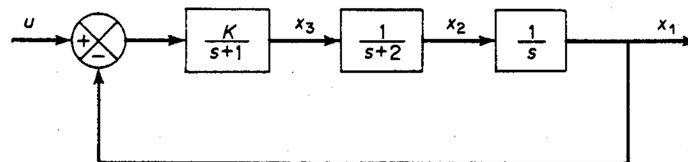


Fig. 15.3 Sistema de controle.

A equação de estado do sistema é

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 1 \\ -K & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ K \end{bmatrix} [u] \tag{15-4}$$

Ao determinar a gama de estabilidade para K, supomos que a entrada u é nula. Então a Eq. (15.4) pode ser escrita como

$$\dot{x}_1 = x_2 \tag{15-5}$$

$$\dot{x}_2 = -2x_2 + x_3 \tag{15-6}$$

$$\dot{x}_3 = -Kx_1 - x_3 \tag{15-7}$$

Das Eqs. (15.5) a (15.7) determinamos que a origem é o estado de equilíbrio. Vamos escolher a matriz **Q** positiva semidefinida real simétrica

$$\mathbf{Q} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{15-8}$$

Esta escolha de **Q** é permitida pois $\dot{V}(\mathbf{x}) = -\mathbf{x}'\mathbf{Q}\mathbf{x}$ não pode ser identicamente zero exceto na origem. Para verificar isto, note que

$$\dot{V}(\mathbf{x}) = -\mathbf{x}'\mathbf{Q}\mathbf{x} = -x_3^2$$

$\dot{V}(\mathbf{x})$, sendo identicamente nulo, implica que x_3 é identicamente zero. Se x_3 é zero, então x_1 deve ser zero, pois, da Eq. (15.7), obtemos

$$0 = -Kx_1 - 0$$

Se x_1 é zero, então x_2 também deve ser zero, pois, da Eq. (15.5),

$$0 = x_2$$

Portanto, $\dot{V}(\mathbf{x})$ é identicamente zero apenas na origem. Deta forma, podemos usar a matriz **Q** definida pela Eq. (15.8) para a análise de estabilidade.

Vamos resolver

$$\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}$$

ou

$$\begin{bmatrix} 0 & 0 & -K \\ 1 & -2 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{bmatrix} +$$

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 1 \\ -K & 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

para os elementos de **P**. O resultado é

$$\mathbf{P} = \begin{bmatrix} \frac{K^2 + 12K}{12 - 2K} & \frac{6K}{12 - 2K} & 0 \\ \frac{6K}{12 - 2K} & \frac{3K}{12 - 2K} & \frac{K}{12 - 2K} \\ 0 & \frac{K}{12 - 2K} & \frac{6}{12 - 2K} \end{bmatrix}$$

Para que **P** seja positiva definida, é necessário e suficiente que

$$12 - 2K > 0 \qquad \text{e} \qquad K > 0$$

ou

$$0 < K < 6$$

Portanto, para $0 < K < 6$, o sistema é estável no sentido convencional, isto é, a origem é assintoticamente estável globalmente.

Análise de estabilidade de Liapunov de sistemas de tempo discreto. Considere o sistema de tempo discreto descrito por

$$\mathbf{x}(k + 1) = \mathbf{G}\mathbf{x}(k) \tag{15-9}$$

onde **x** é um vetor de estado (vetor n-dimensional) e **G** é uma matriz constante $n \times n$. A origem $\mathbf{x} = \mathbf{0}$ é o estado de equilíbrio. Investigaremos a estabilidade deste estado pelo uso do segundo método de Liapunov.

Vamos escolher uma possível função de Liapunov como

$$V(\mathbf{x}(k)) = \mathbf{x}^*(k)\mathbf{P}\mathbf{x}(k)$$

onde **P** é uma matriz Hermitiana positiva definida (ou simétrica real). Observando que para sistemas discretos, ao invés de $\dot{V}(\mathbf{x})$, usamos a diferença entre $V(\mathbf{x}(k + 1))$ e $V(\mathbf{x}(k))$, ou

$$\Delta V(\mathbf{x}(k)) = V(\mathbf{x}(k + 1)) - V(\mathbf{x}(k))$$

que é análogo à derivada de $V(\mathbf{x})$. Então,

$$\begin{aligned} \Delta V(\mathbf{x}(k)) &= V(\mathbf{x}(k + 1)) - V(\mathbf{x}(k)) \\ &= \mathbf{x}^*(k + 1)\mathbf{P}\mathbf{x}(k + 1) - \mathbf{x}^*(k)\mathbf{P}\mathbf{x}(k) \\ &= [\mathbf{G}\mathbf{x}(k)]^*\mathbf{P}[\mathbf{G}\mathbf{x}(k)] - \mathbf{x}^*(k)\mathbf{P}\mathbf{x}(k) \\ &= \mathbf{x}^*(k)\mathbf{G}^*\mathbf{P}\mathbf{G}\mathbf{x}(k) - \mathbf{x}^*(k)\mathbf{P}\mathbf{x}(k) \\ &= \mathbf{x}^*(k)[\mathbf{G}^*\mathbf{P}\mathbf{G} - \mathbf{P}]\mathbf{x}(k) \end{aligned}$$

Como $V(\mathbf{x}(k))$ foi escolhido positivo definido, impomos, para estabilidade assintótica, que $\Delta V(\mathbf{x}(k))$ seja negativo definido. Portanto,

$$\Delta V(\mathbf{x}(k)) = -\mathbf{x}^*(k)\mathbf{Q}\mathbf{x}(k)$$

onde

$$\mathbf{Q} = -(\mathbf{G}^*\mathbf{P}\mathbf{G} - \mathbf{P}) = \text{positiva definida}$$

Portanto, para a estabilidade assintótica do sistema discreto da Eq. (15.9), é suficiente que **Q** seja positiva definida.

De modo similar ao caso de sistemas lineares de tempo contínuo, é conveniente especificar uma matriz positiva definida Hermitiana **Q** (ou simétrica real) e então verificar se a matriz **P** determinada de

$$\mathbf{G}^*\mathbf{P}\mathbf{G} - \mathbf{P} = -\mathbf{Q}$$

é positiva definida. Note que o fato de **P** ser positiva definida é uma condição necessária e suficiente. Resumiremos o que dissemos através de um teorema.

Teorema 15.5 Considere o sistema de tempo discreto

$$\mathbf{x}(k + 1) = \mathbf{G}\mathbf{x}(k)$$

onde **x** é um vetor de estado (vetor n-dimensional) e **G** uma matriz $n \times n$ constante

não singular. Uma condição necessária e suficiente para que o estado de equilíbrio $\mathbf{x} = \mathbf{0}$ seja assintoticamente estável é que, dada uma matriz Hermitiana positiva definida (ou simétrica real) $\mathbf{Q}$, existe uma matriz positiva Hermitiana (ou simétrica real) $\mathbf{P}$ tal que

$$\mathbf{G^*PG - P = -Q}$$

A função escalar $\mathbf{x^*Px}$ é uma função de Liapunov para este sistema.

Se $\Delta V(\mathbf{x}(k)) = -\mathbf{x}^*(k)\mathbf{Q}\mathbf{x}(k)$ não se anula identicamente ao longo de qualquer série solução, então $\mathbf{Q}$ pode ser escolhida positiva semidefinida.

15.5 ESTIMANDO O COMPORTAMENTO DE RESPOSTA TRANSITÓRIA DE SISTEMAS DINÂMICOS

A aplicabilidade do segundo método de Liapunov para teoria de controle não está limitada à análise de estabilidade. Ele pode ser aplicado ao estudo do comportamento de resposta transitória de sistemas lineares e não lineares. Nesta seção discutiremos a relação entre funções de Liapunov e o comportamento de resposta transitória de sistemas dinâmicos.

Como mencionado anteriormente neste capítulo, as funções de Liapunov fornecem uma medida de distância no espaço de estados. Portanto, para um sistema assintoticamente estável para o qual obtivemos uma função de Liapunov, este pode ser usado para estimar a rapidez da resposta transitória.

Vamos supor que o estado de equilíbrio é a origem do espaço de estados e definamos

$$\eta = -\frac{\dot{V}(\mathbf{x}, t)}{V(\mathbf{x}, t)} \tag{15-10}$$

em alguma região do espaço de estados excluindo a origem.

Para sistemas assintoticamente estáveis uniformemente, podemos achar uma função de Liapunov negativa definida $\dot{V}$, pelo menos em princípio. Na análise presente, excluiremos portanto o caso em que $\dot{V}(\mathbf{x}, t) = 0$ para $\mathbf{x} \neq \mathbf{0}$. Então o valor de η é sempre positivo uma vez que V é positivo definido e $\dot{V}$ é negativo definido. A Eq. (15.10) fornece uma indicação de quão rapidamente o sistema se dirige ao seu estado de equilíbrio.

Se fizermos o valor mínimo da relação $-\dot{V}(\mathbf{x}, t)/V(\mathbf{x}, t)$ igual a um certo $\eta_{\text{mín.}}$,

$$\eta_{\text{mín}} = \text{mín}\left[-\frac{\dot{V}(\mathbf{x}, t)}{V(\mathbf{x}, t)}\right] \tag{15-11}$$

então

$$\dot{V}(\mathbf{x}, t) \leq -\eta_{\text{mín}} V(\mathbf{x}, t)$$

e

$$V(\mathbf{x}, t) \leq V(\mathbf{x}_0, t_0)e^{-\eta_{\text{mín}}(t-t_0)} \tag{15-12}$$

onde $V(\mathbf{x}_0, t_0)$ corresponde ao valor inicial de V começando no estado $\mathbf{x} = \mathbf{x}_0$ no instante $t = t_0$.

Interpretando V como a distância a partir da origem, é claro que a Eq. (15.12) fornece uma estimativa de quão rapidamente estamos tendendo ao equilíbrio. Em outras palavras, dado o estado inicial $\mathbf{x}_0$ e o valor inicial $V(\mathbf{x}_0, t_0)$, o sistema alcança a região

$$V(\mathbf{x}, t) = V(\mathbf{x}_0, t_0)e^{-\eta_{\text{mín}}(t-t_0)}$$

em $t - t_0$ segundos.

Da Eq. (15.11), vemos que $1/\eta_{\text{mín.}}$ corresponde à maior constante de tempo relativa a variações na função de Liapunov V. Como $V(\mathbf{x})$ pode ser expandida em uma série começando com um termo quadrático em $\mathbf{x}$, esta constante de tempo $1/\eta_{\text{mín.}}$ é cerca da metade da constante de tempo convencional definida para o sistema. Valores maiores de $\eta_{\text{mín.}}$ correspondem a uma resposta mais rápida.

Como podem existir várias funções de Liapunov para um dado sistema, podemos obter vários diferentes valores de $\eta_{\text{mín.}}$ O maior valor de $\eta_{\text{mín.}}$ representa a figura de mérito. Entretanto, normalmente não sabemos como construir a função de Liapunov que fornece o maior valor de $\eta_{\text{mín.}}$

Para um sistema não linear, $\eta_{\text{mín.}}$ pode depender do estado do sistema. Este método só é útil se a função de Liapunov disponível tem forma suficientemente simples.

Cálculo de η para sistemas lineares invariantes no tempo. Considere o sistema linear invariante no tempo

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$$

onde

$\mathbf{x}$ = vetor n-dimensional
$\mathbf{A}$ = matriz $n \times n$ não singular

Suponha que os autovalores de $\mathbf{A}$ tenham partes reais negativas. Então uma função de Liapunov e sua derivada temporal são

$$V(\mathbf{x}) = \mathbf{x^*Px}, \qquad \dot{V}(\mathbf{x}) = -\mathbf{x^*Qx}$$

onde

$\mathbf{P}$ = matriz positiva definida Hermitiana (ou simétrica real)
$\mathbf{Q} = -(\mathbf{A^*P + PA})$

Uma definição de $\eta_{\text{mín.}}$ equivalente à Eq. (15.11) é

$$\eta_{\text{mín}} = \underset{\mathbf{x}}{\text{mín}}\,\{\mathbf{x^*Qx};\ \mathbf{x^*Px} = 1\}$$

O membro direito desta equação significa que tomamos o valor mínimo de $\mathbf{x^*Qx}$ quando $\mathbf{x}$ varia, satisfazendo a condição de que $\mathbf{x^*Px} = 1$. Isto é o mesmo que considerar o mínimo valor da razão

$$-\frac{\dot{V}(\mathbf{x}, t)}{V(\mathbf{x}, t)} = \frac{\mathbf{x^*Qx}}{\mathbf{x^*Px}}$$

ao longo de uma superfície particular $V(\mathbf{x}) = 1$, ou $\mathbf{x^*Px} = 1$.

A minimização requerida pode ser feita através do uso da técnica de multiplicadores de Lagrange. Seja μ o multiplicador de Lagrange. Então, a minimização de $\mathbf{x^*Qx}$ com a condição de $\mathbf{x^*Px} = 1$ é o mesmo que a minimização de $\mathbf{x^*Qx} - \mu\mathbf{x^*Px}$

com relação a **x**. O valor mínimo de $\mathbf{x}^*\mathbf{Q}\mathbf{x} - \mu\mathbf{x}^*\mathbf{P}\mathbf{x}$ ocorre em um valor $\mathbf{x}_{\text{mín.}}$ de **x** tal que

$$(\mathbf{Q} - \mu\mathbf{P})\mathbf{x}_{\text{mín}} = \mathbf{0} \tag{15-13}$$

Portanto, obtemos

$$\mathbf{x}^*_{\text{mín}}(\mathbf{Q} - \mu\mathbf{P})\mathbf{x}_{\text{mín}} = 0$$

ou

$$\mathbf{x}^*_{\text{mín}}\mathbf{Q}\mathbf{x}_{\text{mín}} = \mu\mathbf{x}^*_{\text{mín}}\mathbf{P}\mathbf{x}_{\text{mín}} = \mu > 0$$

que é um mínimo se μ é mínimo. Mas, pela Eq. (15.13), μ é um autovalor da matriz $\mathbf{QP}^{-1}$. Portanto $\eta_{\text{mín.}}$ é igual ao autovalor mínimo $\lambda_{\text{mín.}}$ de $\mathbf{QP}^{-1}$.

Exemplo 15.7 Considere o seguinte sistema:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Ache uma função de Liapunov para o sistema. Em seguida obtenha um limite superior para o tempo de resposta que o sistema leva para ir de um ponto na borda da curva fechada

$$V(\mathbf{x}) = 150$$

para um ponto dentro da curva fechada

$$V(\mathbf{x}) = 0{,}06$$

Vamos supor que $\mathbf{Q} = \mathbf{I}$ e determinemos **P**, resolvendo

$$\begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

O resultado é

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$$

Portanto obtemos uma função de Liapunov como segue:

$$V = \mathbf{x}'\mathbf{P}\mathbf{x} = \tfrac{1}{2}(3x_1^2 + 2x_1x_2 + 2x_2^2)$$

Observando que

$$\dot{V} = -\mathbf{x}'\mathbf{x} = -(x_1^2 + x_2^2)$$

obtemos o valor de η como

$$\eta = -\frac{\dot{V}}{V} = \frac{2(x_1^2 + x_2^2)}{3x_1^2 + 2x_1x_2 + 2x_2^2}$$

O valor mínimo de η pode ser obtido como o autovalor mínimo de $\mathbf{QP}^{-1}$. Como supusemos que $\mathbf{Q} = \mathbf{I}$, obtemos

$$|\mathbf{QP}^{-1} - \lambda\mathbf{I}| = |\mathbf{P}^{-1} - \lambda\mathbf{I}| = 0$$

A determinação dos autovalores de $\mathbf{P}^{-1}$ equivale a resolver

$$|\mathbf{I} - \mathbf{P}\lambda| = 0$$

ou

$$\begin{vmatrix} 1 - \frac{3}{2}\lambda & -\frac{1}{2}\lambda \\ -\frac{1}{2}\lambda & 1 - \lambda \end{vmatrix} = 0$$

Os dois autovalores são

$$\lambda_1 = 1{,}447, \qquad \lambda_2 = 0{,}533$$

Portanto

$$\eta_{\text{mín}} = \lambda_2 = 0{,}553$$

Da Eq. (15.12),

$$V(\mathbf{x}, t) \leq V(\mathbf{x}_0, t_0)e^{-\eta_{\text{mín}}(t-t_0)}$$

ou

$$-\frac{1}{\eta_{\text{mín}}} \ln\left[\frac{V(\mathbf{x}, t)}{V(\mathbf{x}_0, t_0)}\right] \geq t - t_0 \tag{15-14}$$

Para computar $t - t_0$, usamos o valor de $\eta_{\text{mín.}}$ que acabamos de encontrar. A substituição de $\eta_{\text{mín.}} = 0{,}553$, $V(\mathbf{x}, t) = 0{,}06$, e $V(\mathbf{x}_0, t_0) = 150$ na Eq. (15.14) nos fornece

$$-\frac{1}{0{,}553} \ln\left(\frac{0{,}06}{150}\right) \geq t - t_0$$

ou

$$14{,}1 \geq t - t_0$$

Qualquer trajetória começando na curva $V(\mathbf{x}_0, t_0) = 150$ estará dentro da região circundada pela curva $V(\mathbf{x}, t) = 0{,}06$ dentro de 14,1 unidades de tempo. (Note que este intervalo de tempo depende da escolha das matrizes **P** e **Q**. Se escolhermos matrizes **P** e **Q** apropriadas, poderemos ter um número menor que 14,1 como um limite superior para $t - t_0$.)

15.6 ANÁLISE DE ESTABILIDADE DE SISTEMAS NÃO LINEARES

Em um sistema linear, se o estado de equilíbrio é localmente e assintoticamente estável, então também é assintoticamente estável globalmente. Entretanto, em um sistema não linear, um estado de equilíbrio pode ser assintoticamente estável localmente sem ser assintoticamente estável globalmente. Portanto, as implicações de estabilidade assintótica de estados de equilíbrio de sistemas lineares são bastante diferentes daquelas de sistemas não lineares. Por exemplo, considere o sistema não linear discutido no Exemplo 12.7, reescrito abaixo como

$$\ddot{x} + 0{,}5\dot{x} + 2x + x^2 = 0$$

A Fig. 15.4 mostra o diagrama do plano de fase para este sistema. Neste caso, a origem é assintoticamente estável. Mas trajetórias partindo de pontos fora da região hachurada vão para o infinito, como mostrado. O fato de que a estabilidade assintótica é uma propriedade local pode ser claramente visto neste exemplo.

Como a estabilidade de sistemas não lineares tem uma natureza local, geralmente estamos interessados em determinar uma função de Liapunov que satisfaz condições de estabilidade na maior região em torno da origem.

Vários métodos baseados no segundo método de Liapunov são disponíveis para testar a estabilidade de sistemas não lineares. Esta seção apresenta dois dos métodos: o primeiro é o método de Krasovskii para testar condições suficientes para estabilidade assintótica, e o segundo é o método de Schultz-Gibson, ou do gradiente variável, para gerar funções de Liapunov.

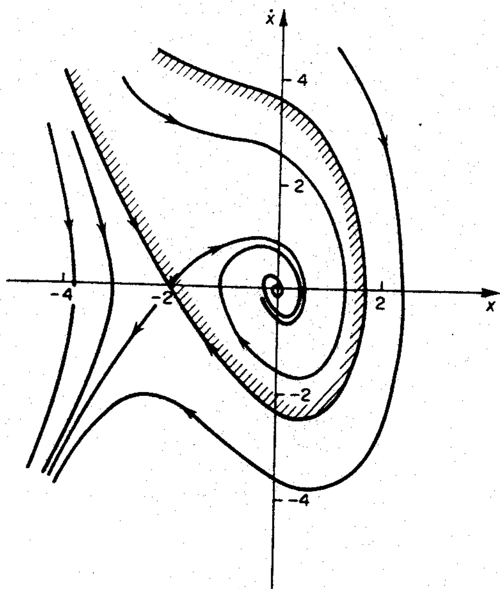


Fig. 15.4 Diagrama de plano de fase do sistema descrito por $\ddot{x} + 0{,}5\dot{x} + 2x + x^2 = 0$.

Método de Krasovskii. Para certos sistemas não lineares, uma possível função de Liapunov pode ser escolhida em termos de $\dot{\mathbf{x}}$ ao invés das variáveis de estado. Krasovskii sugeriu uma possível função de Liapunov como sendo a norma Euclidiana de $\dot{\mathbf{x}}$, ou $V = \|\dot{\mathbf{x}}\|^2$.

O teorema de Krasovskii, a ser apresentado em seguida, não é limitado a pequenos deslocamentos do estado de equilíbrio. Ele difere essencialmente dos métodos de linearização usuais. O teorema de Krasovskii para estabilidade assintótica global fornece condições suficientes para sistemas não lineares e condições necessárias e suficientes* para sistemas lineares. (Um estado de equilíbrio de um sistema não linear pode ser estável mesmo que as condições especificadas neste teorema não sejam satisfeitas. Portanto, ao usar o teorema de Krasovskii, devemos ter cuidado para não tirar conclusões erradas sobre a estabilidade do estado de equilíbrio do dado sistema não linear.)

Em um sistema não linear, pode haver mais de um estado de equilíbrio. Entretanto, é possível transferir um estado de equilíbrio isolado para a origem do espaço de estados por uma transformação adequada de coordenadas. Portanto, consideraremos o estado de equilíbrio em estudo como a origem.

Apresentaremos agora o teorema de Krasovskii.

Teorema 15.6. (Teorema de Krasovskii). Considere o sistema definido por

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

onde $\mathbf{x}$ é um vetor n-dimensional. Suponha que $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ e que $\mathbf{f}(\mathbf{x})$ é diferenciável com respeito a x_i, $i = 1, 2, \ldots n$. A matriz Jacobiana $\mathbf{F}(\mathbf{x})$ para o sistema é

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

Vamos definir

$$\hat{\mathbf{F}}(\mathbf{x}) = \mathbf{F}^*(\mathbf{x}) + \mathbf{F}(\mathbf{x})$$

onde $\mathbf{F}^*(\mathbf{x})$ é a transposta conjugada de $\mathbf{F}(\mathbf{x})$. Se a matriz Hermitiana $\hat{\mathbf{F}}(\mathbf{x})$ é negativa definida, então o estado de equilíbrio $\mathbf{x} = \mathbf{0}$ é assintoticamente estável. Uma função de Liapunov para este sistema é

$$V(\mathbf{x}) = \mathbf{f}^*(\mathbf{x})\mathbf{f}(\mathbf{x})$$

*N. do T.: Mesmo para sistemas lineares as condições são apenas suficientes. Por exemplo, em $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ com $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -4 & -3 \end{bmatrix}$ ou $\mathbf{A} = \begin{bmatrix} -1 & 1 \\ -7 & -3 \end{bmatrix}$, temos $\hat{F}$ não negativa definida.

Se, adicionalmente, $\mathbf{f}^*(\mathbf{x})\mathbf{f}(\mathbf{x}) \to \infty$ quando $\|\mathbf{x}\| \to \infty$, então o estado de equilíbrio é assintoticamente estável globalmente.

DEMONSTRAÇÃO: Se $\hat{\mathbf{F}}(\mathbf{x})$ é negativa definida para todo $\mathbf{x} \neq \mathbf{0}$, o determinante de $\mathbf{F}$ é não nulo em todo espaço exceto em $\mathbf{x} = \mathbf{0}$. (Referir-se aos Problemas A.15.5 e A.15.7.) Não há outro estado de equilíbrio além de $\mathbf{x} = \mathbf{0}$ em todo espaço de estados. Como $\mathbf{f}(\mathbf{0}) = \mathbf{0}$, $\mathbf{f}(\mathbf{x}) \neq \mathbf{0}$ para $\mathbf{x} \neq \mathbf{0}$, e $V(\mathbf{x}) = \mathbf{f}^*(\mathbf{x})\mathbf{f}(\mathbf{x})$, $V(\mathbf{x})$ é positiva definida. Observando que

$$\dot{\mathbf{f}}(\mathbf{x}) = \mathbf{F}(\mathbf{x})\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})\mathbf{f}(\mathbf{x})$$

podemos obter $\dot{V}$ como

$$\begin{aligned}\dot{V}(\mathbf{x}) &= \dot{\mathbf{f}}^*(\mathbf{x})\mathbf{f}(\mathbf{x}) + \mathbf{f}^*(\mathbf{x})\dot{\mathbf{f}}(\mathbf{x})\\ &= [\mathbf{F}(\mathbf{x})\mathbf{f}(\mathbf{x})]^*\mathbf{f}(\mathbf{x}) + \mathbf{f}^*(\mathbf{x})\mathbf{F}(\mathbf{x})\mathbf{f}(\mathbf{x})\\ &= \mathbf{f}^*(\mathbf{x})[\mathbf{F}^*(\mathbf{x}) + \mathbf{F}(\mathbf{x})]\mathbf{f}(\mathbf{x})\\ &= \mathbf{f}^*(\mathbf{x})\hat{\mathbf{F}}(\mathbf{x})\mathbf{f}(\mathbf{x})\end{aligned}$$

Se $\hat{\mathbf{F}}(\mathbf{x})$ é negativa definida, vemos que $\dot{V}(\mathbf{x})$ é negativa definida. Portanto $V(\mathbf{x})$ é uma função de Liapunov. Desta forma, a origem é assintoticamente estável. Se $V(\mathbf{x}) = \mathbf{f}^*(\mathbf{x})\mathbf{f}(\mathbf{x})$ tende a infinito quando $\|\mathbf{x}\| \to \infty$, então, pelo Teorema 15.1, o estado de equilíbrio é assintoticamente estável globalmente.

Note que $\hat{\mathbf{F}}(\mathbf{x})$, sendo negativa definida, requer que $\mathbf{F}(\mathbf{x})$ tenha elementos não nulos em sua diagonal principal. Portanto, se $f_i(\mathbf{x})$ não é função de x_i, então $\hat{\mathbf{F}}(\mathbf{x})$ não pode ser negativa definida.

Exemplo 15.8 Usando o teorema de Krasovskii, examine a estabilidade do estado de equilíbrio $\mathbf{x} = \mathbf{0}$ do seguinte sistema:

$$\begin{aligned}\dot{x}_1 &= -x_1\\ \dot{x}_2 &= x_1 - x_2 - x_2^3\end{aligned}$$

Para este sistema

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} -x_1 \\ x_1 - x_2 - x_2^3 \end{bmatrix}$$

e

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} -1 & 0 \\ 1 & -1 - 3x_2^2 \end{bmatrix}$$

Portanto

$$\begin{aligned}\hat{\mathbf{F}}(\mathbf{x}) &= \mathbf{F}'(\mathbf{x}) + \mathbf{F}(\mathbf{x})\\ &= \begin{bmatrix} -2 & 1 \\ 1 & -2 - 6x_2^2 \end{bmatrix}\end{aligned}$$

Como $\hat{\mathbf{F}}(\mathbf{x})$ é negativa definida para todo $\mathbf{x} \neq \mathbf{0}$, o estado de equilíbrio $\mathbf{x} = \mathbf{0}$ é assintoticamente estável. Além do mais,

$$\mathbf{f}'(\mathbf{x})\mathbf{f}(\mathbf{x}) = x_1^2 + (x_1 - x_2 - x_2^3)^2 \longrightarrow \infty$$

quando $\|\mathbf{x}\| \to \infty$. O estado de equilíbrio $\mathbf{x} = \mathbf{0}$ é, portanto, assintoticamente estável globalmente.

Método do gradiente variável. O método do gradiente variável é baseado no fato de que, se existe uma função particular de Liapunov, que é capaz de provar a estabilidade de um dado sistema, então também existe um gradiente único desta função V.

Considere o sistema descrito pela seguinte equação:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$$

Supomos que o estado de equilíbrio em questão é a origem do espaço de estados. (Como dito antes, sempre podemos transferir qualquer estado de equilíbrio isolado para a origem.) Vamos denotar uma função de Liapunov de tentativa como V. Nesta função, supomos que V é uma função explícita de $\mathbf{x}$ mas não de t. Então,

$$\dot{V} = \frac{\partial V}{\partial x_1}\dot{x}_1 + \frac{\partial V}{\partial x_2}\dot{x}_2 + \cdots + \frac{\partial V}{\partial x_n}\dot{x}_n$$

que pode ser escrita

$$\dot{V} = (\nabla V)'\dot{\mathbf{x}}$$

onde $(\nabla V)'$ é a transposta de $\nabla V \cdot \nabla V$, o gradiente de V, é

$$\nabla V = \begin{bmatrix} \frac{\partial V}{\partial x_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial V}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla V_1 \\ \cdot \\ \cdot \\ \cdot \\ \nabla V_n \end{bmatrix}$$

V é obtido como uma integral de linha de ∇V como

$$V = \int_0^{\mathbf{x}} (\nabla V)' \, d\mathbf{x} \qquad (15\text{-}15)$$

O limite superior de integração nesta equação não implica que V é uma quantidade vetorial, mas sim que a integral é uma integral de linha para um ponto arbitrário $(x_1, x_2, \ldots, x_n)$ no espaço de estados. Esta integral pode ser feita independente do caminho de integração. O caminho mais simples é indicado pela forma expandida da Eq. (15.15):

$$\begin{aligned}V = &\int_0^{x_1, (x_2 = x_3 = \cdots = x_n = 0)} \nabla V_1 \, dx_1 + \int_0^{x_2, (x_1 = x_1, x_3 = x_4 = \cdots = x_n = 0)} \nabla V_2 \, dx_2 + \cdots \\ &+ \int_0^{x_n, (x_1 = x_1, x_2 = x_2, \cdots, x_{n-1} = x_{n-1})} \nabla V_n \, dx_n \qquad (15\text{-}16)\end{aligned}$$

onde ∇V_i é a componente de ∇V na direção x_i.
Vamos definir

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}$$

A integral dada pela Eq. (15.16) diz que o caminho parte da origem e se move ao longo do vetor $\mathbf{e}_1$ para x_1. Deste ponto, o caminho se move na direção do vetor $\mathbf{e}_2$ para x_2. Em seguida, o caminho se move na direção do vetor $\mathbf{e}_3$ para x_3. Desta forma o caminho finalmente alcança o ponto $(x_1, x_2, \ldots, x_n)$.

Para uma função escalar V ser obtida de forma única a partir de uma integral de linha de uma função vetorial ∇V, a matriz $\mathbf{F}$ seguinte, formada tomando-se $\partial \nabla V_i / \partial x_j$,

$$\mathbf{F} = \begin{bmatrix} \frac{\partial \nabla V_1}{\partial x_1} & \frac{\partial \nabla V_1}{\partial x_2} & \cdots & \frac{\partial \nabla V_1}{\partial x_n} \\ \frac{\partial \nabla V_2}{\partial x_1} & \frac{\partial \nabla V_2}{\partial x_2} & \cdots & \frac{\partial \nabla V_2}{\partial x_n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{\partial \nabla V_n}{\partial x_1} & \frac{\partial \nabla V_n}{\partial x_2} & \cdots & \frac{\partial \nabla V_n}{\partial x_n} \end{bmatrix}$$

deve ser simétrica. Portanto, impomos que

$$\frac{\partial \nabla V_i}{\partial x_j} = \frac{\partial \nabla V_j}{\partial x_i} \qquad (i, j = 1, 2, \ldots, n)$$

O número total de tais equações é $n(n-1)/2$. Por exemplo, no caso de $n = 3$, temos três equações:

$$\frac{\partial \nabla V_2}{\partial x_1} = \frac{\partial \nabla V_1}{\partial x_2}, \qquad \frac{\partial \nabla V_3}{\partial x_1} = \frac{\partial \nabla V_1}{\partial x_3}, \qquad \frac{\partial \nabla V_3}{\partial x_2} = \frac{\partial \nabla V_2}{\partial x_3}$$

A condição na matriz $\mathbf{F}$ é portanto uma condição de rotacional generalizado para o caso n-dimensional. O problema de determinar uma função V que satisfaz o teorema de Liapunov é então transformado no problema de determinar um ∇V tal que o rotacional n-dimensional de ∇V é igual a zero. Além do mais, V e $\dot{V}$ determinados a partir de ∇V devem ser suficientes para provar a estabilidade; isto é, devem satisfazer o teorema de Liapunov. Inicialmente fazemos ∇V igual a um vetor coluna arbitrário

$$\nabla V = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdot \\ \cdot \\ \cdot \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{bmatrix} \tag{15-17}$$

Os a_{ij} são quantidades totalmente indeterminadas. Os a_{ij} podem ser constantes ou funções do tempo t e/ou funções das variáveis de estado. Entretanto, é conveniente escolher a_{nn} como uma constante ou uma função do tempo t. Vários dos a_{ij} podem ser escolhidos nulos, ou eles são óbvios a partir das condições em $\dot{V}$ impostas pelo investigador, ou podem ser determinados das equações do rotacional.

Se o estado de equilíbrio $\mathbf{x} = \mathbf{0}$ de um sistema não linear é assintoticamente estável, então poderemos obter uma função de Liapunov seguindo o procedimento:

1. Suponha ∇V da forma da Eq. (15.17).
2. De ∇V, determine $\dot{V}$.
3. Imponha que $\dot{V}$ seja negativa definida ou pelo menos negativa semidefinida.
4. Use as $n(n-1)/2$ equações do rotacional, obtidas da imposição de simetria de F, para determinar os coeficientes desconhecidos restantes em ∇V.
5. Reteste $\dot{V}$, pois a adição de termos requeridos como um resultado do Passo 4 pode alterar $\dot{V}$.
6. Determine V da Eq. (15.16).
7. Examine a região de estabilidade assintótica.

Note que o fato de não conseguir obter uma função de Liapunov adequada através deste método não implica a instabilidade do estado de equilíbrio. (Não se pode tirar conclusões sobre a estabilidade.)

Exemplo 15.9 Vamos construir uma função de Liapunov para o seguinte sistema:

$$\dot{x}_1 = -x_1 + 2x_1^2 x_2$$
$$\dot{x}_2 = -x_2$$

usando o método do gradiente variável.

Seja o gradiente de V

$$\nabla V = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + 2x_2 \end{bmatrix}$$

A derivada de V é então

$$\begin{aligned} \dot{V} &= \nabla V' \dot{\mathbf{x}} \\ &= (a_{11}x_1 + a_{12}x_2)\dot{x}_1 + (a_{21}x_1 + 2x_2)\dot{x}_2 \\ &= -a_{11}x_1^2 + 2a_{11}x_1^3 x_2 - a_{12}x_1x_2 + 2a_{12}x_1^2x_2^2 - a_{21}x_1x_2 - 2x_2^2 \end{aligned} \tag{15-18}$$

Façamos como uma tentativa

$$a_{11} = 1, \qquad a_{12} = a_{21} = 0$$

então a Eq. (15.18) se torna

$$\dot{V} = -x_1^2(1 - 2x_1x_2) - 2x_2^2$$

$\dot{V}$ é negativa definida se

$$1 - 2x_1x_2 > 0 \tag{15-19}$$

Portanto, a Eq. (15.19) é uma condição em x_1 e x_2. O gradiente ∇V é

$$\nabla V = \begin{bmatrix} x_1 \\ 2x_2 \end{bmatrix}$$

Note que

$$\frac{\partial \nabla V_1}{\partial x_2} = \frac{\partial \nabla V_2}{\partial x_1} = 0$$

Vemos que a equação do rotacional é satisfeita. V é

$$V = \int_0^{x_1,(x_2=0)} x_1\,dx_1 + \int_0^{x_2,(x_1=x_1)} 2x_2\,dx_2$$
$$= \frac{x_1^2}{2} + x_2^2 \tag{15-20}$$

Desta função de Liapunov, podemos dizer que o sistema é assintoticamente estável na região

$$1 > 2x_1x_2$$

Para mostrar que a função de Liapunov dada pela Eq. (15.20) não é a única possível, vamos escolher

$$a_{11} = \frac{2}{(1 - x_1x_2)^2}, \qquad a_{12} = \frac{-x_1^2}{(1 - x_1x_2)^2}, \qquad a_{21} = \frac{x_1^2}{(1 - x_1x_2)^2}$$

Então

$$\dot{V} = -2x_1^2 - 2x_2^2$$

$\dot{V}$ é negativa definida em todo o plano de estados. ∇V se torna

$$\nabla V = \begin{bmatrix} \dfrac{2x_1}{(1 - x_1x_2)^2} - \dfrac{x_1^2x_2}{(1 - x_1x_2)^2} \\ \dfrac{x_1^3}{(1 - x_1x_2)^2} + 2x_2 \end{bmatrix}$$

Como

$$\frac{\partial \nabla V_1}{\partial x_2} = \frac{3x_1^2 - x_1^3x_2}{(1 - x_1x_2)^3}$$
$$\frac{\partial \nabla V_2}{\partial x_1} = \frac{3x_1^2 - x_1^3x_2}{(1 - x_1x_2)^3}$$

a equação do rotacional é satisfeita. V é

$$V = \int_0^{x_1,(x_2=0)} \left[\frac{2x_1}{(1 - x_1x_2)^2} - \frac{x_1^2x_2}{(1 - x_1x_2)^2}\right] dx_1 +$$
$$\int_0^{x_2,(x_1=x_1)} \left[\frac{x_1^3}{(1 - x_1x_2)^2} + 2x_2\right] dx_2 = x_2^2 + \frac{x_1^2}{1 - x_1x_2} \tag{15-21}$$

Desta função de Liapunov, vemos que a origem do sistema é assintoticamente estável na região

$$1 > x_1x_2$$

É claro que a função de Liapunov definida pela Eq. (15.20) fornece uma região de estabilidade assintótica menor do que a função de Liapunov dada pela Eq. (15.21). Portanto esta última função é uma escolha melhor do que a anterior.

15.7 COMENTÁRIOS CONCLUSIVOS

Ao concluir este capítulo, resumiremos pontos importantes para serem lembrados:

1. Uma função de Liapunov tem as seguintes propriedades:
   (a) É uma função escalar.
   (b) É uma função positiva definida, pelo menos nas vizinhanças da origem.
   (c) Sua derivada temporal é não-positiva.
   (d) Não é uma função única para um dado sistema.
2. Uma vantagem do segundo método de Liapunov está na sua habilidade de fornecer informação sobre estabilidade global tanto para sistemas lineares quanto não lineares.
3. Se uma função de Liapunov pode ser achada em uma região **Ω** incluindo a origem do espaço de estados, então a estabilidade ou a estabilidade assintótica da origem está verificada. Isto, entretanto, não significa necessariamente que as trajetórias, partindo de um estado da região **Ω**, tendem a infinito, pois o segundo método apenas estabelece condições suficientes para estabilidade. (Lembre-se de que o fato de não conseguir determinar uma função de Liapunov para mostrar estabilidade, estabilidade assintótica, ou instabilidade de um estado de equilíbrio não significa nada.) Note também que para sistemas não lineares certas funções de Liapunov são superiores a outras, significando que a região de estabilidade assintótica é maior, ou $\dot{V}$ é negativa definida ao invés de negativa semidefinida etc.
4. Se a origem do sistema é estável ou assintoticamente estável, então funções de Liapunov com as propriedades requeridas sempre existem, embora a sua determinação possa ser um tanto difícil. Esta dificuldade é a principal desvantagem da análise de estabilidade baseada no segundo método de Liapunov.
5. Além de prover um critério de estabilidade, as funções de Liapunov podem ser usadas para a análise de resposta transitória de sistemas simples. (Veja a Seção 15.5.) As funções de Liapunov também podem ser usadas para resolver problemas de otimização baseados nos índices de desempenho quadráticos. (Veja a Seção 16.5.)

Problema A.15.1 Determine se a seguinte forma quadrática é negativa definida.

$$Q = -x_1^2 - 3x_2^2 - 11x_3^2 + 2x_1x_2 - 4x_2x_3 - 2x_1x_3$$

Solução. A forma quadrática fornecida Q pode ser escrita como

$$Q = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} -1 & 1 & -1 \\ 1 & -3 & -2 \\ -1 & -2 & -11 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Aplicando o critério de Sylvester, achamos

$$-1 < 0, \qquad \begin{vmatrix} -1 & 1 \\ 1 & -3 \end{vmatrix} > 0, \qquad \begin{vmatrix} -1 & 1 & -1 \\ 1 & -3 & -2 \\ -1 & -2 & -11 \end{vmatrix} < 0$$

A forma quadrática é negativa definida.

Problema A.15.2 Considere o sistema

$$\dot{\mathbf{x}} = \mathbf{Ax}$$

onde os autovalores de $\mathbf{A}$ são distintos e têm partes reais negativas. Suponha que a transformação

$$\mathbf{x} = \mathbf{Py}$$

transforma a matriz dos coeficientes na forma diagonal:

$$\dot{\mathbf{y}} = \mathbf{P}^{-1}\mathbf{APy} = \mathbf{Dy}$$

Obtenha uma função de Liapunov para este sistema.

Solução. Considere o seguinte $V(\mathbf{y})$:

$$V(\mathbf{y}) = \mathbf{y}^*\mathbf{y} = \text{positiva definida}$$

A derivada em relação ao tempo de $V(\mathbf{y})$ é

$$\begin{aligned} \dot{V}(\mathbf{y}) &= \dot{\mathbf{y}}^*\mathbf{y} + \mathbf{y}^*\dot{\mathbf{y}} \\ &= \mathbf{y}^*\mathbf{D}^*\mathbf{y} + \mathbf{y}^*\mathbf{Dy} \\ &= 2\mathbf{y}^*\mathbf{Dy} \end{aligned}$$

Como os autovalores de $\mathbf{D}$ e aqueles de $\mathbf{A}$ são idênticos e os autovalores de $\mathbf{A}$ têm partes reais negativas, $\dot{V}(\mathbf{y})$ é negativo definido. Portanto, a função $V(\mathbf{y})$ considerada aqui é uma função de Liapunov.

Problema A.15.3 Considere o sistema de segunda-ordem

$$\dot{\mathbf{x}} = \mathbf{Ax}$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad (a_{ij} = \text{real})$$

Determine a matriz real simétrica $\mathbf{P}$ que satisfaz

$$\mathbf{A'P} + \mathbf{PA} = -\mathbf{I}$$

Em seguida ache a condição para que $\mathbf{P}$ seja positiva definida. (Note que a condição de $\mathbf{P}$ ser positiva definida implica que a origem $\mathbf{x} = \mathbf{0}$ é assintoticamente estável globalmente.)

Solução. A equação

$$\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

fornece as três equações simultâneas seguintes:

$$\begin{aligned} 2(a_{11}p_{11} + a_{21}p_{12}) &= -1 \\ a_{11}p_{12} + a_{21}p_{22} + a_{12}p_{11} + a_{22}p_{12} &= 0 \\ 2(a_{12}p_{12} + a_{22}p_{22}) &= -1 \end{aligned}$$

Resolvendo para obter os p_{ij}, obtemos

$$\mathbf{P} = \frac{1}{2(a_{11} + a_{22})|\mathbf{A}|}\begin{bmatrix} -(|\mathbf{A}| + a_{21}^2 + a_{22}^2) & a_{12}a_{22} + a_{21}a_{11} \\ a_{12}a_{22} + a_{21}a_{11} & -(|\mathbf{A}| + a_{11}^2 + a_{12}^2) \end{bmatrix}$$

$\mathbf{P}$ é positiva definida se

$$p_{11} = -\frac{|\mathbf{A}| + a_{21}^2 + a_{22}^2}{2(a_{11} + a_{22})|\mathbf{A}|} > 0$$

$$|\mathbf{P}| = \frac{(a_{11} + a_{22})^2 + (a_{12} - a_{21})^2}{4(a_{11} + a_{22})^2|\mathbf{A}|} > 0$$

do que obtemos

$$|\mathbf{A}| > 0, \qquad a_{11} + a_{22} < 0$$

como as condições para que $\mathbf{P}$ seja positiva definida.

Problema A.15.4 Considere o movimento de uma nave espacial em torno dos eixos principais de inércia. As equações de Euler são

$$\begin{aligned} A\dot{\omega}_x - (B - C)\omega_y\omega_z &= T_x \\ B\dot{\omega}_y - (C - A)\omega_z\omega_x &= T_y \\ C\dot{\omega}_z - (A - B)\omega_x\omega_y &= T_z \end{aligned}$$

onde A, B e C denotam os momentos de inércia em torno dos eixos principais; ω_x, ω_y e ω_z denotam as velocidades angulares em torno dos eixos principais; e T_x, T_y e T_z são os torques de controle.

Suponha que a nave espacial está com movimentos de rotação anormais em sua órbita. Deseja-se parar estas rotações anormais aplicando torques de controle que se supõe serem

$$T_x = k_1 A\omega_x$$

$$T_y = k_2 B\omega_y$$

$$T_z = k_3 C\omega_z$$

Determine condições suficientes para operação assintoticamente estável do sistema.

Solução. Vamos escolher as variáveis de estado como

$$x_1 = \omega_x, \qquad x_2 = \omega_y, \qquad x_3 = \omega_z$$

Então as equações do sistema se tornam

$$\dot{x}_1 - \left(\frac{B}{A} - \frac{C}{A}\right)x_2x_3 = k_1x_1$$

$$\dot{x}_2 - \left(\frac{C}{B} - \frac{A}{B}\right)x_3x_1 = k_2x_2$$

$$\dot{x}_3 - \left(\frac{A}{C} - \frac{B}{C}\right)x_1x_2 = k_3x_3$$

ou

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} k_1 & \frac{B}{A}x_3 & -\frac{C}{A}x_2 \\ -\frac{A}{B}x_3 & k_2 & \frac{C}{B}x_1 \\ \frac{A}{C}x_2 & -\frac{B}{C}x_1 & k_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

O estado de equilíbrio é a origem, ou $\mathbf{x} = \mathbf{0}$. Se escolhermos

$$\begin{aligned} V(\mathbf{x}) = \mathbf{x}'\mathbf{P}\mathbf{x} = \mathbf{x}' \begin{bmatrix} A^2 & 0 & 0 \\ 0 & B^2 & 0 \\ 0 & 0 & C^2 \end{bmatrix} \mathbf{x} \\ = A^2x_1^2 + B^2x_2^2 + C^2x_3^2 \\ = \text{positiva definida} \end{aligned}$$

então a derivada temporal de $V(\mathbf{x})$ é

$$\dot{V}(\mathbf{x}) = \dot{\mathbf{x}}'\mathbf{P}\mathbf{x} + \mathbf{x}'\mathbf{P}\dot{\mathbf{x}}$$

$$= \mathbf{x}' \begin{bmatrix} k_1 & -\frac{A}{B}x_3 & \frac{A}{C}x_2 \\ \frac{B}{A}x_3 & k_2 & -\frac{B}{C}x_1 \\ -\frac{C}{A}x_2 & \frac{C}{B}x_1 & k_3 \end{bmatrix} \begin{bmatrix} A^2 & 0 & 0 \\ 0 & B^2 & 0 \\ 0 & 0 & C^2 \end{bmatrix} \mathbf{x}$$

$$+ \mathbf{x}' \begin{bmatrix} A^2 & 0 & 0 \\ 0 & B^2 & 0 \\ 0 & 0 & C^2 \end{bmatrix} \begin{bmatrix} k_1 & \frac{B}{A}x_3 & -\frac{C}{A}x_2 \\ -\frac{A}{B}x_3 & k_2 & \frac{C}{B}x_1 \\ \frac{A}{C}x_2 & -\frac{B}{C}x_1 & k_3 \end{bmatrix} \mathbf{x}$$

$$= \mathbf{x}' \begin{bmatrix} 2k_1A^2 & 0 & 0 \\ 0 & 2k_2B^2 & 0 \\ 0 & 0 & 2k_3C^2 \end{bmatrix} \mathbf{x} = -\mathbf{x}'\mathbf{Q}\mathbf{x}$$

Para estabilidade assintótica, a condição suficiente é que $\mathbf{Q}$ seja positiva definida. Portanto, requeremos que

$$k_1 < 0, \qquad k_2 < 0, \qquad k_3 < 0$$

Se os k_i são negativos, então, observando que $V(\mathbf{x}) \to \infty$ quando $\|\mathbf{x}\| \to \infty$, vemos que o estado de equilíbrio é assintoticamente estável globalmente.

Problema A.15.5 Para sistemas físicos, variáveis de estado complexas devem ocorrer em pares conjugados. Considere o sistema

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

Suponha que as variáveis de estado x_1 e x_2 são pares complexos conjugados. Então $\partial f_1/\partial x_1$ é o complexo de $\partial f_2/\partial x_2$ ou

$$\frac{\partial f_1}{\partial x_1} = \overline{\frac{\partial f_2}{\partial x_2}}$$

e também

$$\frac{\partial f_1}{\partial x_2} = \overline{\frac{\partial f_2}{\partial x_1}}$$

Usando este fato, mostre que

$$[\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}]^* = \mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}$$

Supondo que $\mathbf{x}$ é um vetor bidimensional.

Solução

$$\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x} = [\bar{x}_1 \quad \bar{x}_2]\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= [\bar{x}_1 \quad \bar{x}_2]\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}x_1 + \dfrac{\partial f_1}{\partial x_2}x_2 \\ \dfrac{\partial f_2}{\partial x_1}x_1 + \dfrac{\partial f_2}{\partial x_2}x_2 \end{bmatrix}$$

$$= \frac{\partial f_1}{\partial x_1}\bar{x}_1 x_1 + \frac{\partial f_1}{\partial x_2}\bar{x}_1 x_2 + \frac{\partial f_2}{\partial x_1}\bar{x}_2 x_1 + \frac{\partial f_2}{\partial x_2}\bar{x}_2 x_2 \tag{15-22}$$

A transposta conjugada de $\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}$ é

$$[\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}]^* = \overline{\frac{\partial f_1}{\partial x_1}}x_1\bar{x}_1 + \overline{\frac{\partial f_1}{\partial x_2}}x_1\bar{x}_2 + \overline{\frac{\partial f_2}{\partial x_1}}x_2\bar{x}_1 + \overline{\frac{\partial f_2}{\partial x_2}}x_2\bar{x}_2$$

Observando que

$$x_1 = \bar{x}_2, \qquad x_2 = \bar{x}_1$$

$$\frac{\partial f_1}{\partial x_1} = \overline{\frac{\partial f_2}{\partial x_2}}, \qquad \frac{\partial f_1}{\partial x_2} = \overline{\frac{\partial f_2}{\partial x_1}}$$

obtemos

$$[\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}]^* = \frac{\partial f_2}{\partial x_2}\bar{x}_2 x_2 + \frac{\partial f_2}{\partial x_1}\bar{x}_2 x_1 + \frac{\partial f_1}{\partial x_2}\bar{x}_1 x_2 + \frac{\partial f_1}{\partial x_1}\bar{x}_1 x_1 \tag{15-23}$$

Comparando as Eqs (15.22) e (15.23), obtemos

$$[\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}]^* = \mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}$$

Problema A.15.6 Considere o sistema

$$\frac{Y(s)}{U(s)} = \frac{4}{(s+2)(s^2+2s+2)}$$

que pode ser reescrito como

$$\frac{Y(s)}{U(s)} = \frac{2}{s+2}\left[\frac{j}{s+1+j} + \frac{-j}{s+1-j}\right]$$

A representação em diagrama de blocos deste sistema é vista na Fig. 15.5. Vamos escolher as variáveis de estado como mostrado na figura. Supondo que $u = 0$, mostre que

$$[\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}]^* = \mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}$$

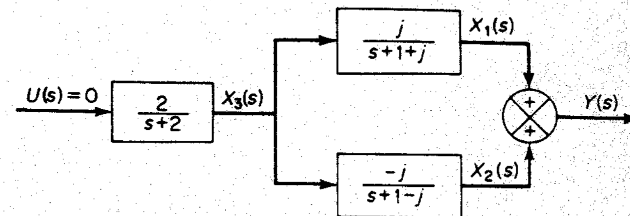


Fig. 15.5 Sistema de controle.

Solução. Do diagrama de blocos, obtemos

$$\frac{X_1}{X_3} = \frac{j}{s+1+j}$$

$$\frac{X_2}{X_3} = \frac{-j}{s+1-j}$$

$$\frac{X_3}{U} = \frac{2}{s+2}$$

ou

$$\dot{x}_1 = -(1+j)x_1 + jx_3$$

$$\dot{x}_2 = -(1-j)x_2 - jx_3$$

$$\dot{x}_3 = -2x_3 + 2u$$

Como u é suposto nulo, obtemos

$$\dot{x}_1 = -(1+j)x_1 + jx_3 = f_1(\mathbf{x})$$

$$\dot{x}_2 = -(1-j)x_2 - jx_3 = f_2(\mathbf{x})$$

$$\dot{x}_3 = -2x_3 = f_3(\mathbf{x})$$

A matriz Jacobiana para este sistema é

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} -(1+j) & 0 & j \\ 0 & -(1-j) & -j \\ 0 & 0 & -2 \end{bmatrix}$$

Observe que

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \bar{x}_2 \\ \bar{x}_1 \\ x_3 \end{bmatrix}$$

(x_3 é uma variável de estado real ou $\bar{x}_3 = x_3$). Agora calcule **x*F(x)x**.

$$\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x} = [\bar{x}_1 \quad \bar{x}_2 \quad x_3]\begin{bmatrix} -(1+j) & 0 & j \\ 0 & -(1-j) & -j \\ 0 & 0 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= [\bar{x}_1 \quad \bar{x}_2 \quad x_3]\begin{bmatrix} -(1+j)x_1 + jx_3 \\ -(1-j)x_2 - jx_3 \\ -2x_3 \end{bmatrix}$$

$$= -(1+j)\bar{x}_1 x_1 + j\bar{x}_1 x_3 - (1-j)\bar{x}_2 x_2 - j\bar{x}_2 x_3 - 2x_3^2 \qquad (15\text{-}24)$$

Agora tome o complexo conjugado de **x*F(x)x**. Da Eq. (15.24),

$$[\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}]^* = -(1-j)x_1\bar{x}_1 - jx_1x_3 - (1+j)x_2\bar{x}_2 + jx_2x_3 - 2x_3^2$$

Observando que $x_1 = \bar{x}_2$, $x_2 = \bar{x}_1$, obtemos

$$[\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}]^* = -(1-j)\bar{x}_2 x_2 - j\bar{x}_2 x_3 - (1+j)\bar{x}_1 x_1 + j\bar{x}_1 x_3 - 2x_3^2 \qquad (15\text{-}25)$$

Comparando as Eqs. (15.24) e (15.25), vemos que

$$[\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}]^* = \mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}$$

Problema A.15.7 Mostre que se

$$\hat{\mathbf{F}}(\mathbf{x}) = \mathbf{F}^*(\mathbf{x}) + \mathbf{F}(\mathbf{x})$$

onde

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

é negativa definida, então o determinante de **F(x)** é não nulo para $\mathbf{x} \neq \mathbf{0}$, ou

$$|\mathbf{F}(\mathbf{x})| \neq 0 \quad \text{para } \mathbf{x} \neq \mathbf{0}$$

Solução. Usando a relação

$$[\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}]^* = \mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}$$

temos

$$\begin{aligned} \mathbf{x}^*\hat{\mathbf{F}}(\mathbf{x})\mathbf{x} &= \mathbf{x}^*[\mathbf{F}^*(\mathbf{x}) + \mathbf{F}(\mathbf{x})]\mathbf{x} \\ &= \mathbf{x}^*\mathbf{F}^*(\mathbf{x})\mathbf{x} + \mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x} \\ &= [\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x}]^* + \mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x} \\ &= 2\mathbf{x}^*\mathbf{F}(\mathbf{x})\mathbf{x} \end{aligned}$$

Como $\hat{\mathbf{F}}(\mathbf{x})$ é negativa definida, **F(x)** também é negativa definida. Portanto,*

$$|\mathbf{F}(\mathbf{x})| \neq 0 \quad \text{para } \mathbf{x} \neq \mathbf{0}^*$$

Problema A.15.8 Determine a estabilidade da origem do sistema linear variante no tempo descrito por

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad \mathbf{A}(t) = \begin{bmatrix} 0 & 1 \\ -\frac{1}{t+1} & -10 \end{bmatrix} \qquad (t \geq 0)$$

Utilize o método do gradiente variável.

Solução. Definamos o gradiente de V como

$$\nabla V = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

Então a derivada de V é

$$\begin{aligned} \dot{V} &= (\nabla V)'\dot{\mathbf{x}} \\ &= (a_{11}x_1 + a_{12}x_2)x_2 + (a_{21}x_1 + a_{22}x_2)\left(-\frac{1}{t+1}x_1 - 10x_2\right) \end{aligned}$$

---

*N. do T.: Como o critério de Sylvester é para matrizes simétricas, esta demonstração pode ser melhor finalizada observando-se as colunas da matriz **F(x)** e notando que são linearmente independentes, e, portanto, o determinante de **F** é não nulo para $\mathbf{x} \neq \mathbf{0}$; ou seja,

$$\mathbf{x}^*\mathbf{F}\mathbf{x} = \mathbf{x}^*[\mathbf{F}\text{col}_1 \ \mathbf{F}\text{col}_2 \ \ldots \mathbf{F}\text{col}_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x}^*[x_1\mathbf{F}\text{col}_1 + x_2\mathbf{F}\text{col}_2 + \ldots + x_n\mathbf{F}\text{col}_n]$$

e como para $\mathbf{x} \neq \mathbf{0}$ esta expressão é não nula (pois $\mathbf{x}^*\hat{\mathbf{F}}^*\mathbf{x} = 2\mathbf{x}^*\mathbf{F}\mathbf{x} \neq 0$ para $\mathbf{x} \neq \mathbf{0}$), as colunas de **F** devem ser linearmente independentes.

Vamos escolher $a_{12} = a_{21} = 0$: então,

$$\dot{V} = a_{11}x_1x_2 + a_{22}x_2\left(-\frac{1}{t+1}x_1 - 10x_2\right)$$

Escolhendo $a_{11} = 1$ e $a_{22} = t + 1$, ∇V pode ser escrito como

$$\nabla V = \begin{bmatrix} x_1 \\ (t+1)x_2 \end{bmatrix}$$

Como

$$\int_0^{x_1,(x_2=0)} x_1\,dx_1 + \int_0^{x_2,(x_1=x_1)} (t+1)x_2\,dx_2 = \frac{1}{2}[x_1^2 + (t+1)x_2^2]$$

Vamos escolher

$$V = \frac{1}{2}[x_1^2 + (t+1)x_2^2]$$

Então

$$\begin{aligned}\dot{V} &= x_1\dot{x}_1 + \frac{x_2^2}{2} + (t+1)x_2\dot{x}_2 \\ &= -(10t + 9{,}5)x_2^2\end{aligned}$$

Portanto $\dot{V}$ é negativa semidefinida. Observando que $\dot{V}$ é identicamente nula apenas na origem $x_1 = x_2 = 0$, pelo Teorema 15.2, concluímos que o estado de equilíbrio na origem $\mathbf{x} = \mathbf{0}$ é assintoticamente estável globalmente.

PROBLEMAS

Problema B.15.1 Determine se as seguintes formas quadráticas são positivas definidas.

$$Q = x_1^2 + 4x_2^2 + x_3^2 + 2x_1x_2 - 6x_2x_3 - 2x_1x_3$$

Problema B.15.2 Determine uma função de Liapunov para o seguinte sistema:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 2 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Problema B.15.3 Determine a estabilidade do estado de equilíbrio do seguinte sistema:

$$x_1(k+1) = x_1(k) + 3x_2(k)$$

$$x_2(k+1) = -3x_1(k) - 2x_2(k) - 3x_3(k)$$

$$x_3(k+1) = x_1(k)$$

Problema B.15.4 Determine a estabilidade da origem do seguinte sistema:

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -g(x_1) + x_2\end{aligned}$$

onde $g(x_1)/x_1 > 0$ para $x_1 \neq 0$.

Problema B.15.5 Considere o sistema descrito por

$$\begin{aligned}\dot{x}_1 &= -g_1(x_1) + g_2(x_1, x_2) \\ \dot{x}_2 &= g_3(x_2)\end{aligned}$$

onde $g_1(0) = g_3(0) = 0$, $g_2(0, x_2) = 0$. Obtenha condições suficientes para a estabilidade da origem do sistema.

Problema B.15.6 Determine a estabilidade da origem do seguinte sistema:

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1^3 - x_2\end{aligned}$$

Problema B.15.7 Determine a estabilidade da origem do seguinte sistema:

$$\begin{aligned}\dot{x}_1 &= -x_1 + x_2 + x_1(x_1^2 + x_2^2) \\ \dot{x}_2 &= -x_1 - x_2 + x_2(x_1^2 + x_2^2)\end{aligned}$$

Problema B.15.8 Determine a estabilidade da origem do seguinte sistema:

$$\begin{aligned}\dot{x}_1 &= x_1 - x_2 - x_1^3 \\ \dot{x}_2 &= x_1 + x_2 - x_2^3\end{aligned}$$

16

Sistemas de Controle Ótimo e Adaptativos

16.1 INTRODUÇÃO

Sistemas de controle ótimo. Problemas de controle ótimo receberam grande atenção durante a década passada devido à crescente demanda de sistemas de grande desempenho e à pronta disponibilidade do computador digital.

O conceito de otimização de sistemas de controle engloba uma seleção de um índice de desempenho e de um projeto que fornece o sistema de controle ótimo dentro de limites impostos por condições limitantes de ordem física. Tal sistema de controle ótimo difere de um sistema ideal, em que o primeiro é o melhor que se pode conseguir na presença de limitações físicas, enquanto que o segundo pode bem ser um objetivo inalcançável.

Índices de desempenho. Ao resolver problemas de sistemas de controle ótimo, podemos ter o objetivo de achar uma regra para a determinação da decisão de controle presente, sujeita a certas limitações, que minimizará alguma medida de um desvio do comportamento ideal. Tal medida é normalmente provida por um critério de otimização, ou índice de desempenho. O índice de desempenho, que foi definido no Cap. 7, é uma função cujo valor indica quão bem o desempenho atual do sistema se compara com o desempenho desejado. Na maioria dos casos particulares, o comportamento do sistema é otimizado escolhendo-se o vetor de controle de tal forma que o índice de desempenho é minimizado (ou maximizado).

O índice de desempenho é importante porque determina, em grande parte, a natureza do controle ótimo resultante. Isto é, o controle resultante pode ser linear, não-linear, estacionário ou variante no tempo, dependendo da forma do índice de desempenho. O engenheiro de controle formula este índice baseado nos requerimentos do problema. Portanto, ele influencia a natureza do sistema resultante. Os requisitos do problema normalmente incluem não apenas requisitos de desempenho mas também restrições na forma do controle para assegurar realizabilidade física.

O processo de otimização deve prover não apenas leis de controle, configurações de parâmetros que são ótimas, mas também uma medida da degradação no desempenho pelo afastamento da função índice de desempenho do seu valor mínimo (ou máximo) que resulta do uso de leis de controle não ótimas.

Em um grau considerável, o uso da teoria de otimização no projeto de sistemas é dificultado pelo conflito entre tratabilidade analítica e utilidade prática na seleção do índice de desempenho. É desejável que os critérios para controle ótimo se originem não de um ponto de vista matemático mas de um ponto de vista de aplicação. Entretanto, em geral, a escolha de um índice de desempenho envolve um compromisso entre uma avaliação significativa do desempenho do sistema e um problema matemático tratável.

A escolha do índice de desempenho mais apropriado para um dado problema é muito difícil, especialmente em sistemas complicados. Por exemplo, considere o problema da maximização da carga útil de uma nave espacial. A carga útil pode ser considerada como a diferença entre o peso do veículo após terminada a missão e os componentes residuais do veículo, tais como estruturas de suporte, equipamentos de comunicações, de potência e de controle de altitude. A maximização da carga útil, portanto, envolverá uma otimização tanto do programa de propulsão e projeto da missão para mínimos gastos de combustível, como um projeto ótimo dos componentes do veículo. Em aplicações espaciais, outras especificações de desempenho podem ser o gasto mínimo de combustível, erro mínimo de atingir um alvo, tempo mínimo etc. Em aplicações civis de controle, as considerações principais normalmente são de ordem econômica.

Formulação de problemas de otimização. As quantidades que aparecem em problemas de otimização de sistemas de controle são variáveis de estado, variáveis de controle, e parâmetros de sistema. Considere, por exemplo, um veículo que é assimilado a uma massa pontual que se locomove no espaço. As variáveis de estado deste sistema podem ser as três coordenadas de posição do veículo, as três coordenadas de velocidade, e a massa instantânea do veículo. Estas variáveis de estado são geradas de um conjunto de equações diferenciais, que, neste exemplo, podem ser simplesmente as equações de Newton para o movimento e uma equação de continuidade relacionando o fluxo de combustível com a taxa de perda de massa do veículo. As variáveis de controle, para este exemplo, podem ser a magnitude da força vertical do veículo e um conjunto de ângulos que definem a direção da força. Os parâmetros do sistema são as constantes que descrevem certas propriedades do problema. Tais parâmetros podem ser a velocidade de exaustão do sistema de propulsão ou um tempo pré-especificado de término da força vertical. Para sistemas que utilizam propulsão iônica, estes parâmetros podem ser os valores da velocidade de exaustão e o tamanho da unidade geradora de potência a bordo do veículo.

Em geral, o problema de otimização de sistemas de controle pode ser formulado se as seguintes informações são dadas:

1. equações de estado e de saída
2. vetor de controle
3. restrições do problema
4. índice de desempenho
5. parâmetros do sistema

Um problema de controle ótimo é a determinação do vetor de controle ótimo $\mathbf{u}(t)$ dentro da classe de vetores de controle admissíveis. Este vetor $\mathbf{u}(t)$ normalmente depende de

1. estado inicial ou saída inicial
2. estado desejado ou saída desejada
3. natureza das restrições
4. natureza do índice de desempenho

Exceto para casos especiais, o problema pode ser tão complicado para solução analítica que se deve obter uma solução por computador.

Neste capítulo, discutiremos sistemas de controle de tempo ótimo e sistemas de controle baseados em índices de desempenho quadráticos.

Sistemas de controle de tempo ótimo. Na Seção 16.4, discutiremos dois problemas de controle de tempo ótimo. Um é, dado o sistema de tempo discreto,

$$\mathbf{x}((k+1)T) = \mathbf{G}(T)\mathbf{x}(kT) + \mathbf{H}(T)\mathbf{u}(kT)$$

onde a norma de $\mathbf{u}(kT)$ não é limitada, determinar o vetor de controle $\mathbf{u}(kT)$ que trará qualquer estado inicial para a origem do espaço de estado no número mínimo de períodos de amostragem.

O outro problema de controle de tempo ótimo é, dado o processo descrito por

$$\overset{(n)}{y} + a_1 \overset{(n-1)}{y} + \cdots + a_{n-1}\dot{y} + a_n y = b_0 \overset{(n)}{u} + b_1 \overset{(n-1)}{u} + \cdots + b_{n-1}\dot{u} + b_n u$$

determinar a função de controle u que transferirá a saída y do sistema de uma dada condição inicial para zero em um tempo mínimo T. A função de controle u deve satisfazer uma restrição de magnitude e a saída e suas derivadas devem permanecer nulas para tempo $t \geqslant T$ se não se aplicam subseqüentes perturbações.

A obtenção da solução explícita para o controle de tempo ótimo de um processo de ordem n é extremamente difícil. Portanto, discutiremos apenas um caso simples onde o processo é descrito por

$$\ddot{y} + a_1\dot{y} + a_2 y = bu$$

Isto é, a função de transferência do processo apresenta dois pólos e nenhum zero.

Sistemas de controle ótimo baseados em índices de desempenho quadráticos. Em muitos sistemas de controle, na prática, queremos minimizar alguma função do sinal de erro. Por exemplo, dado o sistema

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

podemos desejar minimizar uma função de erro generalizado tal como

$$J = \int_0^T [\boldsymbol{\xi}(t) - \mathbf{x}(t)]^*\mathbf{Q}[\boldsymbol{\xi}(t) - \mathbf{x}(t)]\, dt$$

onde $\boldsymbol{\xi}(t)$ representa o estado desejado, $\mathbf{x}(t)$ o estado atual [portanto, $\boldsymbol{\xi}(t) - \mathbf{x}(t)$ é o vetor de erro], $\mathbf{Q}$ uma matriz positiva-definida (ou positiva semidefinida), e o intervalo de tempo $0 \leqslant t \leqslant T$ ou é finito ou infinito.

Além de considerar erros como uma medida do desempenho do sistema, entretanto, devemos normalmente dar atenção à energia requerida para a ação de controle. Como $u(t)$ pode ter dimensões de força ou torque, a energia de controle é proporcional à integral de $[u(t)]^2$. Se os erros são minimizados sem levar em conta a energia requerida, então pode resultar um projeto que requer valores extremamente grandes de $u(t)$. Isto é indesejável uma vez que todos os sistema físicos estão sujeitos a saturação. Portanto, considerações práticas impõem uma restrição ao vetor de controle; por exemplo,

$$\int_0^T \mathbf{u}^*(t)\mathbf{R}\mathbf{u}(t)\, dt = K$$

onde $\mathbf{R}$ é uma matriz positiva-definida e K é uma constante positiva. O índice de desempenho de um sistema de controle no intervalo de tempo $0 \leqslant t \leqslant T$ pode ser escrito, usando-se um multiplicador de Lagrange λ, como segue:

$$J = \int_0^T [\boldsymbol{\xi}(t) - \mathbf{x}(t)]^*\mathbf{Q}[\boldsymbol{\xi}(t) - \mathbf{x}(t)]\, dt + \lambda \int_0^T \mathbf{u}^*(t)\mathbf{R}\mathbf{u}(t)\, dt \quad (0 \leq t \leq T) \tag{16.1}$$

O multiplicador de Lagrange λ é uma constante positiva indicando o peso do custo de controle com relação aos erros minimizantes.

A determinação da lei de controle ótimo para o sistema,

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

sujeita ao índice de desempenho dado pela Eq. (16.1), tem um significado prático de que o sistema resultante estabelece um compromisso entre a minimização do erro integral e a minimização da energia de controle. Na Seção 16.5, consideramos um caso especial onde $T = \infty$, o estado desejado $\boldsymbol{\xi}$ é a origem, ou $\boldsymbol{\xi} = \mathbf{0}$, e o vetor de estado é real. Nestas condições, o índice de desempenho quadrático dado pela Eq. (16.1) pode ser expresso como

$$J = \int_0^\infty [\mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}'(t)\mathbf{R}\mathbf{u}(t)]\, dt$$

onde incluímos λ na matriz positiva definida $\mathbf{R}$. Note que neste problema $\mathbf{u}(t)$ não tem restrições. Veremos adiante que tal índice de desempenho quadrático resulta em operações matemáticas convenientes na solução de problemas de controle ótimo.

Comentários sobre sistemas de controle ótimo. O sistema que minimiza (ou maximiza) o índice de desempenho selecionado é, por definição, ótimo. É evidente que o índice de desempenho na realidade determina a configuração do sistema. É importante ressaltar que um sistema de controle ótimo sob um certo índice de desempenho é, em geral, não ótimo sob outros índices de desempenho. Além do mais, a realização física de uma particular lei de controle ótimo pode ser um tanto difícil e cara. Portanto pode não ter sentido fazer grandes gastos em implementar um controlador ótimo que é o melhor em algum sentido restrito ou individual. Um sistema de controle raramente é projetado para realizar uma única tarefa especificada totalmente *a priori*. Ao invés disto, ele é projetado para realizar uma tarefa selecionada ao acaso de um repertório completo de possíveis tarefas. Em sistemas práticos, pode então ser mais aconselhável procurar leis de controle ótimo aproximadas que não estão rigidamente ligadas com um índice de desempenho único.

Devemos notar que um sistema de controle ótimo, obtido matematicamente, nos dá, na maioria das aplicações práticas, a limitação de desempenho fundamental sob o índice de desempenho dado e serve mais como uma medida do que como um objetivo prático. Portanto, antes de decidir implementar o sistema de controle ótimo ou algo inferior, porém mais simples, devemos avaliar com cuidado uma medida do grau em que o desempenho do sistema de controle ótimo complexo excede o do sistema mais simples subótimo. A não ser que o sistema de controle ótimo possa ser justificado, não construiremos sistemas de controle ótimo extremamente complicados.

Uma vez que a limitação de desempenho fundamental é determinada usando-se a teoria de controle ótimo, devemos fazer esforços para projetar um sistema simples que está perto do ótimo. Mantendo isto em mente, construímos um protó-

tipo do sistema físico, o testamos e modificamos até que seja obtido um sistema satisfatório, que tem características de desempenho perto das do sistema de controle ótimo obtido usando-se a teoria de controle ótimo.

Perguntas relativas à existência da solução dos problemas de controle ótimo. Como já dissemos, o problema de controle ótimo, dado qualquer estado inicial $\mathbf{x}(t_0)$, é o de determinar um vetor de controle admissível $\mathbf{u}(t)$ que transfere o estado para a região desejada do espaço de estados e para o qual o índice de desempenho é minimizado.

É importante mencionar que, em alguns casos, uma combinação particular de processo, estado desejado, índice de desempenho, e restrições pode tornar um controle ótimo impossível. Isto se relaciona com o desejo de obter um desempenho além das capacidades físicas do sistema.

Perguntas relativas à existência de um sistema de controle ótimo são importantes uma vez que servem para informar ao projetista se um dado índice de desempenho é ou não realista para um dado sistema e conjunto de restrições. Duas das perguntas mais importantes são as de controlabilidade e observabilidade. Em seguida, explicaremos brevemente o que significa controlabilidade e observabilidade.

Controlabilidade e observabilidade. Um sistema é dito controlável no instante t_0 se é possível, usando-se um vetor de controle arbitrário (sem restrições), transferir o sistema de qualquer estado inicial $\mathbf{x}(t_0)$ para qualquer outro estado em um intervalo finito de tempo.

Um sistema é dito observável no instante t_0 se, com o sistema no estado $\mathbf{x}(t_0)$, é possível determinar este estado a partir da observação da saída durante um intervalo de tempo finito.

Os conceitos de controlabilidade e observabilidade foram introduzidos por Kalman. Eles têm papel importante no controle ótimo de sistemas multivariáveis. De fato, as condições de controlabilidade e observabilidade podem definir a existência de uma solução completa para o problema de controle ótimo.

Resumo do capítulo. Este capítulo encaminha o leitor a alguns dos sistemas de controle ótimo e adaptativo mais importantes. A Seção 16.2 apresenta as condições para controlabilidade completa e a Seção 16.3 nos dá as condições para observabilidade completa. A Seção 16.4 discute o controle de tempo ótimo de sistemas de tempo discreto e de tempo contínuo.

Os sistemas de controle ótimo baseados em índices de desempenho quadrático são apresentados na Seção 16.5. Aqui é apresentado o método de Liapunov. A Seção 16.6 discute brevemente sistemas de controle de referência-modelo. As duas seções finais, Seções 16.7 e 16.8, apresentam discussões introdutórias sobre sistemas de controle adaptativos. Tais sistemas medem continuamente o desempenho do sistema e ajustam seus parâmetros para manter um desempenho quase-ótimo.

16.2 CONTROLABILIDADE

A solução para um problema de controle ótimo pode não existir se o sistema considerado não é controlável. Embora a maioria dos sistemas físicos sejam controláveis e observáveis, os correspondentes modelos matemáticos podem não ter a propriedade de controlabilidade e observabilidade. Então é necessário saber as condições em que um sistema é controlável e observável.

Nesta seção, primeiro definiremos a independência linear de vetores e então derivaremos as condições para a controlabilidade completa de estado. Iniciaremos nossa discussão com sistemas de tempo discreto uma vez que são mais simples de serem analisados que os de tempo contínuo. Então derivaremos as condições para controlabilidade completa de estado para sistemas de tempo contínuo. Finalmente, discutiremos a controlabilidade completa de saída.

Independência linear de vetores. Os vetores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ são ditos linearmente independentes se

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n = \mathbf{0}$$

onde $c_1, c_2, \ldots, c_n$ são constantes, implica

$$c_1 = c_2 = \cdots = c_n = 0$$

Por outro lado, os vetores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ são ditos linearmente dependentes se e somente se $\mathbf{x}_i$ podem ser representados como uma combinação linear de $\mathbf{x}_j$ ($j = 1, 2, \ldots, n; j \neq i$), ou

$$\mathbf{x}_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} c_j \mathbf{x}_j$$

para algum conjunto de constantes c_j. Isto significa que se $\mathbf{x}_i$ pode ser expresso como uma combinação linear dos outros vetores do conjunto, ele é linearmente dependente deles ou não é um membro independente do conjunto.

Exemplo 16.1 Os vetores

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$$

são linearmente dependentes uma vez que

$$\mathbf{x}_1 + \mathbf{x}_2 - \mathbf{x}_3 = \mathbf{0}$$

Os vetores

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{y}_3 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

são linearmente independentes uma vez que

$$c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + c_3\mathbf{y}_3 = \mathbf{0}$$

implica

$$c_1 = c_2 = c_3 = 0$$

Note que se uma matriz $n \times n$ é não-singular (isto é, a matriz tem característica n ou o determinante é não nulo), então n colunas (ou linhas) são linearmente independentes. Se a matriz $n \times n$ é singular (isto é, a característica da matriz é menor do que n ou o determinante é zero), então n colunas (ou linhas) são linearmente dependentes. Para demonstrar isto, note que

$$[\mathbf{x}_1 \mid \mathbf{x}_2 \mid \mathbf{x}_3] = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 0 & 2 \\ 3 & 1 & 4 \end{bmatrix} = \text{singular}$$

$$[\mathbf{y}_1 \mid \mathbf{y}_2 \mid \mathbf{y}_3] = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 0 & 2 \\ 3 & 1 & 2 \end{bmatrix} = \text{não singular}$$

Controlabilidade completa de estado de sistemas de tempo discreto. Considere o sistema de tempo discreto descrito por

$$\mathbf{x}((k+1)T) = \mathbf{G}\mathbf{x}(kT) - \mathbf{H}u(kT) \tag{16.2}$$

onde

$\mathbf{x}(kT)$ = vetor de estado (vetor n-dimensional)
$u(kT)$ = sinal de controle
$\mathbf{G}$ = matriz $n \times n$
$\mathbf{H}$ = matriz $n \times 1$
T = período de amostragem

Note que $u(kT)$ é constante para $kT \leqslant t < (k-1)T$. Sem perda de generalidade, podemos supor que o estado inicial é arbitrário e que o estado final é a origem do espaço de estados.

O sistema de tempo discreto dado pela Eq. (16.2) é de estado controlável se existe um sinal de controle constante por trechos $u(kT)$ definido sobre um intervalo finito de amostragem $0 \leqslant kT < nT$ tal que, começando de qualquer estado inicial, o estado $\mathbf{x}(kT)$ pode ser feito nulo para $kT \geqslant nT$. Se todo estado é controlável, então o sistema é dito de estado completamente controlável.

Derivaremos a condição para a controlabilidade completa de estado usando o fato de que, se um sistema é de estado completamente controlável, então existe um sinal de controle contínuo por trechos que transferirá qualquer estado inicial para a origem em um número finito de períodos de amostragem.

A solução da Eq. (16.2) é

$$\mathbf{x}(kT) = \mathbf{G}^k\mathbf{x}(0) + \sum_{j=0}^{k-1} \mathbf{G}^{k-j-1}\mathbf{H}u(jT)$$

Se o sistema é controlável, então, partindo de um $\mathbf{x}(0)$ arbitrário, podemos levar o estado para a origem, ou $\mathbf{x}(kT) = \mathbf{0}$ para $k \geqslant n$ pela aplicação de algum $u(0), u(T), \ldots, u((n-1)T)$. Portanto,

$$\mathbf{G}^n\mathbf{x}(0) + \sum_{j=0}^{n-1} \mathbf{G}^{n-j-1}\mathbf{H}u(jT) = \mathbf{0}$$

ou

$$\begin{aligned} \mathbf{x}(0) &= -\sum_{j=0}^{n-1} \mathbf{G}^{-(j+1)}\mathbf{H}u(jT) \\ &= -[\mathbf{G}^{-1}\mathbf{H}u(0) + \mathbf{G}^{-2}\mathbf{H}u(T) + \cdots + \mathbf{G}^{-n}\mathbf{H}u((n-1)T)] \end{aligned} \tag{16.3}$$

Como $\mathbf{G}$ é uma matriz $n \times n$ não-singular e $\mathbf{H}$ uma matriz $n \times 1$, claramente $\mathbf{G}^{-1}\mathbf{H}$, $\mathbf{G}^{-2}\mathbf{H}, \ldots, \mathbf{G}^{-n}\mathbf{H}$ são matrizes $n \times 1$ ou vetores-coluna. Vamos definir

$$\mathbf{G}^{-1}\mathbf{H} = \begin{bmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{n1} \end{bmatrix}, \quad \mathbf{G}^{-2}\mathbf{H} = \begin{bmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{n2} \end{bmatrix}, \quad \ldots, \quad \mathbf{G}^{-n}\mathbf{H} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{nn} \end{bmatrix}$$

Então a Eq. (16.3) pode ser reescrita como as seguintes n equações algébricas simultâneas:

$$\begin{aligned} x_1(0) &= -a_{11}u(0) - a_{12}u(T) - \cdots - a_{1n}u((n-1)T) \\ x_2(0) &= -a_{21}u(0) - a_{22}u(T) - \cdots - a_{2n}u((n-1)T) \\ &\cdots \\ x_n(0) &= -a_{n1}u(0) - a_{n2}u(T) - \cdots - a_{nn}u((n-1)T) \end{aligned}$$

Se estas equações devem fornecer uma solução para qualquer conjunto de valores $x_1(0), x_2(0), \ldots, x_n(0)$, então imporemos que a matriz $n \times n$ (a_{ij}) seja não-singular. (Em caso contrário a solução pode não existir.) A condição de que a matriz $n \times n$ (a_{ij}) seja não-singular implica que cada coluna da matriz (a_{ij}) é linearmente independente, ou a matriz (a_{ij}) tem característica n.

Da análise anterior, podemos enunciar as condições para a controlabilidade completa de estado do sistema da Eq. (16.2) como segue: O sistema é de estado completamente controlável se e apenas se os vetores $\mathbf{G}^{-1}\mathbf{H}, \mathbf{G}^{-2}\mathbf{H}, \ldots, \mathbf{G}^{-n}\mathbf{H}$ são linearmente independentes ou a matriz $n \times n$

$$[\mathbf{G}^{-1}\mathbf{H} \mid \mathbf{G}^{-2}\mathbf{H} \mid \cdots \mid \mathbf{G}^{-n}\mathbf{H}]$$

tem característica n. Como $\mathbf{G}$ é não-singular, esta condição pode ser formulada como a característica da matriz

$$[\mathbf{H} \mid \mathbf{G}^{-1}\mathbf{H} \mid \cdots \mid \mathbf{G}^{-(n-1)}\mathbf{H}]$$

é n. Esta condição também pode ser formulada com a característica da matriz

$$[\mathbf{H} \mid \mathbf{G}\mathbf{H} \mid \cdots \mid \mathbf{G}^{n-1}\mathbf{H}]$$

sendo n.

Note que se o sistema da Eq. (16.2) é de estado completamente controlável, então podemos transferir qualquer estado inicial para a origem no máximo em n períodos de amostragem. (Lembrar que isto é verdade se e apenas se a magnitude de $u(kT)$ é não-limitada. Se a magnitude de $u(kT)$ é limitada, o controle para a origem

pode tomar mais do que n períodos de amostragem.)

Se o sistema é descrito por

$$\mathbf{x}((k+1)T) = \mathbf{G}\mathbf{x}(kT) + \mathbf{H}\mathbf{u}(kT)$$

onde $\mathbf{u}(kT)$ é um vetor de dimensão r, então pode-se provar que a condição para controlabilidade completa de estado é que a seguinte matriz $n \times nr$

$$[\mathbf{G}^{-1}\mathbf{H} \mid \mathbf{G}^{-2}\mathbf{H} \mid \cdots \mid \mathbf{G}^{-n}\mathbf{H}]$$

tem característica n. Esta condição é a mesma que a condição de a característica da matriz

$$[\mathbf{H} \mid \mathbf{G}^{-1}\mathbf{H} \mid \cdots \mid \mathbf{G}^{-(n-1)}\mathbf{H}]$$

ser n, ou a característica da matriz

$$[\mathbf{H} \mid \mathbf{G}\mathbf{H} \mid \cdots \mid \mathbf{G}^{n-1}\mathbf{H}]$$

ser n.

Controlabilidade completa de estado de sistema de tempo contínuo. Considere o sistema de tempo contínuo

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \tag{16.4}$$

onde

$\mathbf{x}$ = vetor de estado (vetor de dimensão n)
u = sinal de controle
$\mathbf{A}$ = matriz $n \times n$
$\mathbf{B}$ = matriz $n \times 1$

O sistema descrito pela Eq. (16.4) é dito de estado controlável em $t = t_0$ se é possível construir um sinal de controle não-limitado que transferirá um estado inicial para qualquer estado final em um intervalo de tempo finito $t_0 \leq t \leq t_1$. Se todo estado é controlável, então o sistema é dito de estado completamente controlável.

Derivaremos agora a condição para controlabilidade completa de estado. Sem perda de generalidade, podemos supor que o estado final é a origem do espaço de estados e que o tempo inicial é zero, ou $t_0 = 0$.

A solução da Eq. (16.4) é

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}u(\tau)\,d\tau$$

Aplicando a definição de controlabilidade completa de estado que acabamos de dar, temos

$$\mathbf{x}(t_1) = \mathbf{0} = e^{\mathbf{A}t_1}\mathbf{x}(0) + \int_0^{t_1} e^{\mathbf{A}(t_1-\tau)}\mathbf{B}u(\tau)\,d\tau$$

ou

$$\mathbf{x}(0) = -\int_0^{t_1} e^{-\mathbf{A}\tau}\mathbf{B}u(\tau)\,d\tau \tag{16.5}$$

Note que $e^{-\mathbf{A}\tau}$ pode ser escrita como

$$e^{-\mathbf{A}\tau} = \sum_{k=0}^{n-1} \alpha_k(\tau)\mathbf{A}^k \tag{16.6}$$

Substituindo a Eq. (16.6) na Eq. (16.5), temos

$$\mathbf{x}(0) = -\sum_{k=0}^{n-1} \mathbf{A}^k\mathbf{B} \int_0^{t_1} \alpha_k(\tau)u(\tau)\,d\tau \tag{16.7}$$

Vamos fazer

$$\int_0^{t_1} \alpha_k(\tau)u(\tau)\,d\tau = \beta_k$$

Então a Eq. (16.7) se torna

$$\mathbf{x}(0) = -\sum_{k=0}^{n-1} \mathbf{A}^k\mathbf{B}\beta_k$$

$$= -[\mathbf{B} \mid \mathbf{AB} \mid \cdots \mid \mathbf{A}^{n-1}\mathbf{B}] \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{n-1} \end{bmatrix} \tag{16.8}$$

Se o sistema é de estado completamente controlável, então, dado qualquer estado inicial $\mathbf{x}(0)$, a Eq. (16.8) deve ser satisfeita. Isto requer que a característica da matriz $n \times n$

$$[\mathbf{B} \mid \mathbf{AB} \mid \cdots \mid \mathbf{A}^{n-1}\mathbf{B}]$$

seja n.

Desta análise, podemos formular a condição para controlabilidade completa de estado como segue: O sistema dado pela Eq. (16.4) é de estado completamente controlável se e somente se os vetores $\mathbf{B}$, $\mathbf{AB}$, ..., $\mathbf{A}^{n-1}\mathbf{B}$ são linearmente independentes ou a matriz $n \times n$

$$[\mathbf{B} \mid \mathbf{AB} \mid \cdots \mid \mathbf{A}^{n-1}\mathbf{B}]$$

tem característica n.

O resultado que acabamos de obter pode ser estendido para o caso onde o vetor de controle $\mathbf{u}$ é r-dimensional. Se o sistema é descrito por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

onde $\mathbf{u}$ é um vetor r-dimensional, então pode-se provar que a condição para controlabilidade completa de estado é que a seguinte matriz $n \times nr$

$$[\mathbf{B} \mid \mathbf{AB} \mid \cdots \mid \mathbf{A}^{n-1}\mathbf{B}]$$

deve ter característica n, ou que ela contém n vetores-coluna linearmente independentes.

Exemplo 16.2 Considere o sistema dado por

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}1 & 1\\ 0 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}1\\ 0\end{bmatrix}[u]$$

Como

$$[\mathbf{B} \mid \mathbf{AB}] = \begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix} = \text{singular}$$

o sistema não é de estado completamente controlável.

Exemplo 16.3 Considere o sistema dado por

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}1 & 1\\ 2 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix}[u]$$

Para este caso,

$$[\mathbf{B} \mid \mathbf{AB}] = \begin{bmatrix}0 & 1\\ 1 & -1\end{bmatrix} = \text{não singular}$$

O sistema é portanto de estado completamente controlável.

Forma alternativa da condição para controlabilidade completa de estado. Considere o sistema definido por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{16.9}$$

onde

$\mathbf{x}$ = vetor de estado (vetor n-dimensional)
$\mathbf{u}$ = vetor de controle (vetor r-dimensional)
$\mathbf{A}$ = matriz $n \times n$
$\mathbf{B}$ = matriz $n \times r$

Se os autovalores de **A** são distintos, então é possível achar uma matriz de transformação **P** tal que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{bmatrix}\lambda_1 & & & & 0\\ & \lambda_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & \lambda_n\end{bmatrix}$$

Note que se os autovalores de **A** são distintos, então os autovetores de **A** também o são,† entretanto, a inversa é verdadeira. Por exemplo, uma matriz real simétrica $n \times n$ com autovalores múltiplos tem n autovetores distintos. Note também que cada coluna da matriz **P** é um autovetor de **A** associado com λ_i ($i = 1, 2, \ldots, n$).

Vamos definir

$$\mathbf{x} = \mathbf{P}\mathbf{z} \tag{16.10}$$

Substituindo a Eq. (16.10) na Eq. (16.9), obtemos

$$\dot{\mathbf{z}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{z} + \mathbf{P}^{-1}\mathbf{B}\mathbf{u} \tag{16.11}$$

Definindo

$$\mathbf{P}^{-1}\mathbf{B} = \mathbf{F} = (f_{ij})$$

podemos reescrever a Eq. (16.11) como

$$\begin{aligned}\dot{z}_1 &= \lambda_1 z_1 + f_{11}u_1 + f_{12}u_2 + \cdots + f_{1r}u_r\\ \dot{z}_2 &= \lambda_2 z_2 + f_{21}u_1 + f_{22}u_2 + \cdots + f_{2r}u_r\\ &\cdots\\ \dot{z}_n &= \lambda_n z_n + f_{n1}u_1 + f_{n2}u_2 + \cdots + f_{nr}u_r\end{aligned}$$

Se os elementos de qualquer linha da matriz $n \times r$ **F** são todos nulos, então a correspondente variável de estado não pode ser controlada por nenhum dos u_i. Portanto, a condição de controlabilidade completa de estado é que se os autovetores de **A** são distintos, então o sistema é de estado completamente controlável se e apenas se nenhuma linha de $\mathbf{P}^{-1}\mathbf{B}$ tem elementos todos nulos. É importante notar que para aplicar esta condição para controlabilidade completa de estado, devemos colocar a matriz $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ na Eq. (16.11) na forma diagonal.

Se a matriz **A** na Eq. (16.9) não tem autovetores distintos, então a diagonalização é impossível. Em tal caso, podemos transformar **A** na forma canônica de Jordan. Se, por exemplo, **A** tem autovalores $\lambda_1, \lambda_1, \lambda_1, \lambda_4, \lambda_4, \lambda_6, \ldots, \lambda_n$ e tem $n - 3$ autovetores distintos, então a forma canônica de Jordan de **A** é

$$\mathbf{J} = \left[\begin{array}{ccc|cc|cccc}\lambda_1 & 1 & 0 & & & & & & 0\\ 0 & \lambda_1 & 1 & & & & & & \\ 0 & 0 & \lambda_1 & & & & & & \\ \hline & & & \lambda_4 & 1 & & & & \\ & & & & \lambda_4 & & & & \\ \hline & & & & & \lambda_6 & & & \\ & & & & & & \cdot & & \\ & & & & & & & \cdot & \\ 0 & & & & & & & & \lambda_n\end{array}\right]$$

†N.T.: Para maior clareza, onde se ler *autovetores distintos* de **A**, deve-se entender que **A** tem n autovetores (não generalizados) linearmente independentes.

As submatrizes 3 × 3 e 2 × 2 na diagonal principal são chamadas de blocos de Jordan.

Suponha que podemos achar uma matriz de transformação **S** tal que

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{J}$$

Se definimos o novo vetor de estado **z** por

$$\mathbf{x} = \mathbf{S}\mathbf{z} \tag{16.12}$$

então a substituição da Eq. (16.12) na Eq. (16.9) nos dá

$$\begin{aligned}\dot{\mathbf{z}} &= \mathbf{S}^{-1}\mathbf{A}\mathbf{S}\mathbf{z} + \mathbf{S}^{-1}\mathbf{B}\mathbf{u} \\ &= \mathbf{J}\mathbf{z} + \mathbf{S}^{-1}\mathbf{B}\mathbf{u}\end{aligned} \tag{16.13}$$

A condição para controlabilidade completa de estado do sistema da Eq. (16.9) pode ser formulada como segue: O sistema é de estado completamente controlável se e apenas se: (1) não há dois blocos de Jordan em **J** da Eq. (16.13) associados com os mesmos autovalores, (2) os elementos da qualquer linha de $\mathbf{S}^{-1}\mathbf{B}$ que correspondem à última linha de cada bloco de Jordan não são todos nulos, e (3) os elementos de cada linha de $\mathbf{S}^{-1}\mathbf{B}$ que correspondem a autovalores distintos não são todos nulos.

Exemplo 16.4 Os seguintes sistemas são de estado completamente controlável:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}-1 & 0\\ 0 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}2\\ 5\end{bmatrix}[u]$$

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} = \begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 0\\ 0 & 0 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} + \begin{bmatrix}0\\ 4\\ 3\end{bmatrix}[u]$$

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\\ \dot{x}_5\end{bmatrix} = \left[\begin{array}{ccc|cc}-2 & 1 & 0 & & 0\\ 0 & -2 & 1 & & \\ 0 & 0 & -2 & & \\ \hline & & & -5 & 1\\ 0 & & & 0 & -5\end{array}\right]\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\\ x_5\end{bmatrix} + \begin{bmatrix}0 & 1\\ 0 & 0\\ 3 & 0\\ 0 & 0\\ 2 & 1\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$

Os seguintes sistemas não são de estado completamente controlável:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}-1 & 0\\ 0 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}2\\ 0\end{bmatrix}[u]$$

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} = \begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 0\\ 0 & 0 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} + \begin{bmatrix}4 & 2\\ 0 & 0\\ 3 & 0\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\\ \dot{x}_5\end{bmatrix} = \left[\begin{array}{ccc|cc}-2 & 1 & 0 & & 0\\ 0 & -2 & 1 & & \\ 0 & 0 & -2 & & \\ \hline & & & -5 & 1\\ 0 & & & 0 & -5\end{array}\right]\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\\ x_5\end{bmatrix} + \begin{bmatrix}4\\ 2\\ 1\\ 3\\ 0\end{bmatrix}[u]$$

Condição para controlabilidade completa de estado no plano *s*. A condição para controlabilidade completa de estado pode ser formulada em termos de funções de transferência ou matrizes de transferência.

Uma condição necessária e suficiente para controlabilidade completa de estado é que não haja cancelamento na função de transferência ou matriz de transferência. Se houver cancelamento, o sistema não pode ser controlado na direção do modo cancelado.

Exemplo 16.5 Considere a seguinte função de transferência:

$$\frac{X(s)}{U(s)} = \frac{s+2.5}{(s+2.5)(s-1)}$$

Claramente ocorre cancelamento do fator $(s+2{,}5)$ no numerador e denominador desta função de transferência. (Portanto um grau de liberdade é perdido.) Em virtude deste cancelamento, este sistema não é de estado completamente controlável.

A mesma conclusão pode ser obtida escrevendo-se esta função de transferência na forma de equações de estado. Uma representação por espaço de estados é

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}0 & 1\\ 2{,}5 & -1{,}5\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}1\\ 1\end{bmatrix}[u]$$

Como

$$[\mathbf{B} \mid \mathbf{AB}] = \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$$

a característica da matriz $[\mathbf{B} \mid \mathbf{AB}]$ é um. Portanto, chegamos à mesma conclusão: O sistema não é de estado completamente controlável.

Controlabilidade de saída. No projeto prático de um sistema de controle, podemos desejar controlar a saída e não o estado do sistema. A controlabilidade completa de estado não é uma condição necessária nem suficiente para controlar a saída do sistema. Por esta razão, é desejável definir separadamente a controlabilidade completa de saída.

Considere o sistema descrito por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{16.14}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \tag{16.15}$$

onde

x = vetor de estado (vetor n-dimensional)

$\mathbf{u}$ = vetor de controle (vetor r-dimensional)
$\mathbf{y}$ = vetor de saída (vetor m-dimensional)
$\mathbf{A}$ = matriz $n \times n$
$\mathbf{B}$ = matriz $n \times r$
$\mathbf{C}$ = matriz $m \times n$
$\mathbf{D}$ = matriz $m \times r$

O sistema descrito pelas Eqs. (16.14) e (16.15) é dito de saída completamente controlável se é possível construir um vetor de controle não-limitado $\mathbf{u}(t)$ que transferirá qualquer saída inicial $\mathbf{y}(t_0)$ para qualquer saída $\mathbf{y}(t_1)$ em um intervalo de tempo finito $t_0 \leqslant t \leqslant t_1$.

Pode-se provar que a condição para controlabilidade completa de saída é como segue: O sistema descrito pelas Eqs. (16.14) e (16.15) é de saída completamente controlável se e apenas se a matriz $m \times (n + 1)r$

$$[\mathbf{CB} \mid \mathbf{CAB} \mid \mathbf{CA}^2\mathbf{B} \mid \cdots \mid \mathbf{CA}^{n-1}\mathbf{B} \mid \mathbf{D}]$$

tem característica igual a m.

16.3 OBSERVABILIDADE

Nesta seção discutimos a observabilidade de sistemas lineares. Considere o sistema não excitado descrito pelas seguintes equações:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{16.16}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} \tag{16.17}$$

onde

$\mathbf{x}$ = vetor de estado (vetor n-dimensional)
$\mathbf{y}$ = vetor de saída (vetor m-dimensional)
$\mathbf{A}$ = matriz $n \times n$
$\mathbf{C}$ = matriz $m \times n$

O sistema é dito ser completamente observável se todo estado inicial $\mathbf{x}(0)$ pode ser determinado a partir da observação de $\mathbf{y}(t)$ durante um intervalo de tempo finito. O sistema portanto é completamente observável se toda transição do estado afeta cada elemento do vetor de saída. O conceito de observabilidade é útil para resolver o problema de reconstruir variáveis de estado não mensuráveis a partir das mensuráveis no mínimo espaço de tempo possível. (Note que, na prática, a dificuldade encontrada com sistemas de controle ótimo é que algumas das variáveis de estado não são acessíveis para medição direta. Então é necessário estimar as variáveis de estado não mensuráveis para construir os sinais de controle ótimo.)

A razão pela qual consideramos o sistema não excitado é a seguinte: Se o sistema é descrito por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$
$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

então

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\, d\tau$$

e $\mathbf{y}(t)$ é

$$\mathbf{y}(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{x}(0) + \mathbf{C}\int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\, d\tau$$

Como as matrizes $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$ são conhecidas e $\mathbf{u}(t)$ também é conhecida, o termo em integral no lado direito desta última equação é uma grandeza conhecida. Portanto, ela pode ser subtraída do valor observado de $\mathbf{y}(t)$. Portanto, para investigar uma condição necessária e suficiente para observabilidade completa, é suficiente considerar o sistema descrito pelas Eqs. (16.16) e (16.17).

Observabilidade completa de sistemas de tempo discreto. Considere o sistema descrito por

$$\mathbf{x}((k + 1)T) = \mathbf{G}\mathbf{x}(kT) \tag{16.18}$$

$$\mathbf{y}(kT) = \mathbf{C}\mathbf{x}(kT) \tag{16.19}$$

onde

$\mathbf{x}(kT)$ = vetor de estado (vetor n-dimensional)
$\mathbf{y}(kT)$ = vetor de saída (vetor m-dimensional)
$\mathbf{G}$ = matriz $n \times n$
$\mathbf{C}$ = matriz $m \times n$
T = período de amostragem

Este sistema é completamente observável se, dada a saída $\mathbf{y}(kT)$ para períodos de amostragem finitos, é possível determinar o vetor de estado inicial $\mathbf{x}(0)$.

A seguir, derivaremos a condição para observabilidade completa do sistema de tempo discreto descrito pelas Eqs. (16.18) e (16.19). Como a solução $\mathbf{x}(kT)$ é

$$\mathbf{x}(kT) = \mathbf{G}^k\mathbf{x}(0)$$

obtemos

$$\mathbf{y}(kT) = \mathbf{C}\mathbf{G}^k\mathbf{x}(0)$$

A observabilidade completa significa que, dados $\mathbf{y}(0)$, $\mathbf{y}(T)$,..., $\mathbf{y}(NT)$, podemos determinar $x_1(0), x_2(0), \ldots, x_n(0)$. Para determinar n incógnitas, necessitamos apenas de n valores de $\mathbf{y}(kT)$. Portanto, $N = n - 1$. Para um sistema completamente observável, dado

$$\mathbf{y}(0) = \mathbf{C}\mathbf{x}(0)$$
$$\mathbf{y}(T) = \mathbf{C}\mathbf{G}\mathbf{x}(0)$$
$$\cdots$$
$$\mathbf{y}((n-1)T) = \mathbf{C}\mathbf{G}^{n-1}\mathbf{x}(0)$$

devemos poder determinar $x_1(0), x_2(0), \ldots, x_n(0)$. Observando que $\mathbf{y}$ é um vetor de dimensão m, as n equações simultâneas anteriores fornecem nm equações, todas

envolvendo $x_1(0), x_2(0), \ldots, x_n(0)$. Para obter um conjunto único de soluções $x_1(0), x_2(0), \ldots, x_n(0)$ destas nm equações, devemos ser capazes de obter exatamente n equações linearmente independentes entre elas. Isto requer que a matriz $nm \times n$

$$\begin{bmatrix} \mathbf{C} \\ \hdashline \mathbf{CG} \\ \hdashline \cdot \\ \cdot \\ \cdot \\ \hdashline \mathbf{CG}^{n-1} \end{bmatrix}$$

tenha característica n.

Notando que a característica de uma matriz e aquela da transposta conjugada da matriz são iguais, podemos formular a condição para observabilidade completa como segue: O sistema descrito pelas Eqs. (16.18) e (16.19) é completamente observável se e apenas se a matriz $n \times nm$

$$[\mathbf{C}^* \mid \mathbf{G}^*\mathbf{C}^* \mid \cdots \mid (\mathbf{G}^*)^{n-1}\mathbf{C}^*]$$

tem característica n, ou tem n vetores-coluna linearmente independentes.

Observabilidade completa de sistemas de tempo contínuo. Considere o sistema descrito pelas Eqs. (16.16) e (16.17), reescritas como

$$\dot{\mathbf{x}} = \mathbf{Ax}$$
$$\mathbf{y} = \mathbf{Cx}$$

O vetor de saída $\mathbf{y}(t)$ é

$$\mathbf{y}(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{x}(0)$$

Observando que

$$e^{\mathbf{A}t} = \sum_{k=0}^{n-1} \alpha_k(t)\mathbf{A}^k$$

obtemos

$$\mathbf{y}(t) = \sum_{k=0}^{n-1} \alpha_k(t)\mathbf{CA}^k\mathbf{x}(0)$$

ou

$$\mathbf{y}(t) = \alpha_0(t)\mathbf{Cx}(0) + \alpha_1(t)\mathbf{CAx}(0) + \cdots + \alpha_{n-1}(t)\mathbf{CA}^{n-1}\mathbf{x}(0) \qquad (16.20)$$

Se o sistema é completamente observável, então, dada uma saída $\mathbf{y}(t)$ durante um intervalo de tempo $0 \leq t \leq t_1$, $\mathbf{x}(0)$ é unicamente determinado a partir da Eq. (16.20). Pode-se mostrar que isto requer que a característica da matriz $nm \times n$

$$\begin{bmatrix} \mathbf{C} \\ \hdashline \mathbf{CA} \\ \hdashline \cdot \\ \cdot \\ \cdot \\ \hdashline \mathbf{CA}^{n-1} \end{bmatrix}$$

seja n.

Desta análise, podemos formular a condição para observabilidade completa como segue: O sistema descrito pelas Eqs. (16.16) e (16.17) é completamente observável se e apenas se a matriz $n \times nm$

$$[\mathbf{C}^* \mid \mathbf{A}^*\mathbf{C}^* \mid \cdots \mid (\mathbf{A}^*)^{n-1}\mathbf{C}^*]$$

tem característica n, ou tem n vetores-coluna linearmente independentes.

Exemplo 16.6 Considere o sistema descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [u]$$

$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Este sistema é controlável e observável?

Como a característica da matriz

$$[\mathbf{B} \mid \mathbf{AB}] = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$$

é dois, o sistema é de estado completamente controlável.

Para controlabilidade de saída, vamos achar a característica da matriz $[\mathbf{CB} \mid \mathbf{CAB}]$. Como

$$[\mathbf{CB} \mid \mathbf{CAB}] = [0 \quad 1]$$

a característica desta matriz é um. Portanto o sistema é de saída completamente controlável.

Para testar a condição de observabilidade, examine a característica de $[\mathbf{C}' \mid \mathbf{A}'\mathbf{C}']$. Como

$$[\mathbf{C}' \mid \mathbf{A}'\mathbf{C}'] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

a característica de $[\mathbf{C}' \mid \mathbf{A}'\mathbf{C}']$ é dois. Portanto o sistema é completamente observável.

Condições para observabilidade completa no plano s. As condições para observabilidade completa podem também ser formuladas em termos de funções de transferência ou matrizes de transferência. A condição necessária e suficiente para observabilidade completa é que não haja cancelamento na função de transferência ou matriz de transferência. Se houver cancelamento, o modo cancelado não pode ser observado na saída.

Exemplo 16.7 Mostre que o sistema seguinte não é completamente observável:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u$$
$$y = \mathbf{C}\mathbf{x}$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{C} = [4 \quad 5 \quad 1]$$

Note que a função de controle u não afeta a observabilidade completa do sistema. Para examinar a observabilidade completa, podemos simplesmente fazer $u = 0$. Para este sistema, temos

$$[\mathbf{C}' \mid \mathbf{A}'\mathbf{C}' \mid (\mathbf{A}')^2\mathbf{C}'] = \begin{bmatrix} 4 & -6 & 6 \\ 5 & -7 & 5 \\ 1 & -1 & -1 \end{bmatrix}$$

Note que

$$\begin{vmatrix} 4 & -6 & 6 \\ 5 & -7 & 5 \\ 1 & -1 & -1 \end{vmatrix} = 0$$

Portanto a característica da matriz $[\mathbf{C}' \mid \mathbf{A}'\mathbf{C}' \mid (\mathbf{A}')^2\mathbf{C}']$ é menor do que três. Desta forma, o sistema não é completamente observável.

De fato, neste sistema, ocorre cancelamento na função de transferência do sistema. A função de transferência entre $X_1(s)$ e $U(s)$ é

$$\frac{X_1(s)}{U(s)} = \frac{1}{(s+1)(s+2)(s+3)}$$

e a função de transferência entre $Y(s)$ e $X_1(s)$ é

$$\frac{Y(s)}{X_1(s)} = (s+1)(s+4)$$

Portanto a função de transferência entre a saída $Y(s)$ e a entrada $U(s)$ é

$$\frac{Y(s)}{U(s)} = \frac{(s+1)(s+4)}{(s+1)(s+2)(s+3)}$$

Claramente, os dois fatores $(s + 1)$ se cancelam. Isto significa que há estados iniciais $\mathbf{x}(0)$ que não podem ser determinados a partir de medições de $y(t)$.

Relações entre controlabilidade, observabilidade e funções de transferência. A função de transferência não tem cancelamento se e apenas se o sistema é de estado completamente controlável e completamente observável. Isto significa que a função de transferência cancelada não contém toda a informação que caracteriza o sistema dinâmico.

Forma alternativa da condição de observabilidade completa. Considere o sistema descrito pelas Eqs. (16.16) e (16.17), reescritas como

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{16.21}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} \tag{16.22}$$

Suponha que a matriz de transformação **P** transforma **A** para a matriz diagonal, ou

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$$

onde **D** é uma matriz diagonal. Vamos definir

$$\mathbf{x} = \mathbf{P}\mathbf{z}$$

Então as Eqs. (16.21) e (16.22) podem ser escritas como

$$\dot{\mathbf{z}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{z} = \mathbf{D}\mathbf{z}$$
$$\mathbf{y} = \mathbf{C}\mathbf{P}\mathbf{z}$$

Portanto,

$$\mathbf{y}(t) = \mathbf{C}\mathbf{P}e^{\mathbf{D}t}\mathbf{z}(0)$$

ou

$$\mathbf{y}(t) = \mathbf{C}\mathbf{P}\begin{bmatrix} e^{\lambda_1 t} & & & 0 \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_n t} \end{bmatrix}\mathbf{z}(0) = \mathbf{C}\mathbf{P}\begin{bmatrix} e^{\lambda_1 t}z_1(0) \\ e^{\lambda_2 t}z_2(0) \\ \vdots \\ e^{\lambda_n t}z_n(0) \end{bmatrix}$$

O sistema é completamente observável se nenhuma das colunas da matriz $m \times n$ **CP** consiste em elementos todos nulos. Isto acontece porque, se a i-ésima coluna de **CP** consiste em todos os elementos nulos, então a variável de estado $z_i(0)$ não aparecerá na equação de saída e portanto não pode ser determinada a partir da observação de $\mathbf{y}(t)$. Assim, $\mathbf{x}(0)$, que está relacionado com $\mathbf{z}(0)$ através da matriz não-singular **P**, não pode ser determinada. (Lembrar que este teste se aplica apenas se a matriz $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ está na forma diagonal.)

Se a matriz **A** não pode ser transformada na forma diagonal, então, usando uma matriz de transformação adequada **S**, podemos transformar **A** para a forma canônica de Jordan, ou

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{J}$$

onde **J** está na forma canônica de Jordan.

Vamos definir

$$\mathbf{x} = \mathbf{S}\mathbf{z}$$

Então as Eqs. (16.21) e (16.22) podem ser escritas como

$$\dot{\mathbf{z}} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}\mathbf{z} = \mathbf{J}\mathbf{z}$$
$$\mathbf{y} = \mathbf{C}\mathbf{S}\mathbf{z}$$

Portanto

$$\mathbf{y}(t) = \mathbf{C}\mathbf{S}e^{\mathbf{J}t}\mathbf{z}(0)$$

O sistema é completamente observável se: (1) não há dois blocos de Jordan associados com os mesmos autovalores, (2) não há colunas de **CS** que correspondem à primeira linha de cada bloco de Jordan contendo elementos todos nulos, e (3) não há colunas de **CS** que correspondem a autovalores distintos com elementos todos nulos.

Para esclarecer a condição (2) acima, no Exemplo 16.8, enquadramos com linhas pontilhadas as colunas de **CS** que correspondem à primeira linha de cada bloco de Jordan.

Exemplo 16.8 Os seguintes sistemas são completamente observáveis:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad y = [1 \quad 3] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 & & 0 \\ 0 & 2 & 1 & & \\ 0 & 0 & 2 & & \\ & & & -3 & 1 \\ 0 & & & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}, \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

Os seguintes sistemas não são completamente observáveis:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad y = [0 \quad 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 3 \\ 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 & & 0 \\ 0 & 2 & 1 & & \\ 0 & 0 & 2 & & \\ & & & -3 & 1 \\ 0 & & & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}, \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

Princípio da dualidade. Discutiremos agora a relação entre controlabilidade e observabilidade. Introduziremos o princípio da dualidade, devido a Kalman, para esclarecer as analogias aparentes entre controlabilidade e observabilidade.

Considere o sistema S_1 descrito por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$
$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

onde

$\mathbf{x}$ = vetor de estado (vetor n-dimensional)
$\mathbf{u}$ = vetor de controle (vetor r-dimensional)
$\mathbf{y}$ = vetor de saída (vetor m-dimensional)
$\mathbf{A}$ = matriz $n \times n$
$\mathbf{B}$ = matriz $n \times r$
$\mathbf{C}$ = matriz $m \times n$

e o sistema dual S_2 definido por

$$\dot{\mathbf{z}} = \mathbf{A}^*\mathbf{z} + \mathbf{C}^*\mathbf{v}$$
$$\mathbf{n} = \mathbf{B}^*\mathbf{z}$$

onde

$\mathbf{z}$ = vetor de estado (vetor n-dimensional)
$\mathbf{v}$ = vetor de controle (vetor m-dimensional)
$\mathbf{n}$ = vetor de saída (vetor r-dimensional)
$\mathbf{A}^*$ = transposta conjugada de $\mathbf{A}$
$\mathbf{B}^*$ = transposta conjugada de $\mathbf{B}$
$\mathbf{C}^*$ = transposta conjugada de $\mathbf{C}$

O princípio da dualidade diz que o sistema S_1 é de estado completamente controlável (observável) se e apenas se o sistema S_2 é completamente observável (de estado controlável).

Para verificar este princípio, vamos escrever as condições necessárias e suficientes para controlabilidade completa de estado e observabilidade completa dos sistemas S_1 e S_2.

Para o sistema S_1:

1. Uma condição necessária e suficiente para controlabilidade completa de estado é que a característica da matriz $n \times nr$

$$[\mathbf{B} \mid \mathbf{A}\mathbf{B} \mid \cdots \mid \mathbf{A}^{n-1}\mathbf{B}]$$

seja n.

2. Uma condição necessária e suficiente para observabilidade completa é que a característica da matriz $n \times nm$

$$[\mathbf{C}^* \mid \mathbf{A}^*\mathbf{C}^* \mid \cdots \mid (\mathbf{A}^*)^{n-1}\mathbf{C}^*]$$

seja n.

Para o sistema S_2:

1. Uma condição necessária e suficiente para controlabilidade completa de estado é que a característica da matriz $n \times nm$

$$[\mathbf{C}^* \mid \mathbf{A}^*\mathbf{C}^* \mid \cdots \mid (\mathbf{A}^*)^{n-1}\mathbf{C}^*]$$

seja n.

2. Uma condição necessária e suficiente para observabilidade completa é que a característica da matriz $n \times nr$

$$[\mathbf{B} \mid \mathbf{AB} \mid \cdots \mid \mathbf{A}^{n-1}\mathbf{B}]$$

seja n.

Comparando estas condições, fica aparente que este princípio é verdadeiro. Este princípio, naturalmente, se aplica igualmente para sistemas de tempo discreto. Usando este princípio, a observabilidade de um dado sistema pode ser testada verificando-se a controlabilidade de estado de seu dual.

16.4 SISTEMAS DE CONTROLE DE TEMPO ÓTIMO

Nesta seção, consideraremos primeiro o controle de tempo ótimo de sistemas de tempo discreto e, em seguida, o controle de tempo ótimo de sistemas de tempo contínuo.

Controle de tempo ótimo de sistemas de tempo discreto. Se o sistema de ordem n é de tempo discreto, então um estado inicial arbitrário pode ser levado a um estado desejado no máximo em n períodos de amostragem se o sistema é de estado completamente controlável e a norma do vetor de controle é não-limitada.

Considere o seguinte sistema de tempo discreto:

$$\mathbf{x}((k+1)T) = \mathbf{G}(T)\mathbf{x}(kT) + \mathbf{H}(T)\mathbf{u}(kT), \qquad \|\mathbf{u}(kT)\| \leq \infty$$

onde

$\mathbf{x}$ = vetor de estado (vetor n-dimensional)
$\mathbf{u}$ = vetor de controle (vetor r-dimensional)
$\mathbf{G}(T)$ = matriz constante $n \times n$
$\mathbf{H}(T)$ = matriz constante $n \times r$

(Note que o período de amostragem T é uma constante.) Se este sistema é de estado completamente controlável, então podemos transferir um estado inicial arbitrário $\mathbf{x}(0)$ para a origem em n passos, ou

$$\mathbf{x}(nT) = \mathbf{G}^n(T)\mathbf{x}(0) + \sum_{k=0}^{n-1} \mathbf{G}^{n-k-1}(T)\mathbf{H}(T)\mathbf{u}(kT) = \mathbf{0}$$

O vetor de controle desejado $\mathbf{u}(kT)$ pode ser determinado resolvendo-se as seguintes n equações matriciais simultâneas:

$$\begin{aligned}
\mathbf{x}(T) &= \mathbf{G}(T)\mathbf{x}(0) + \mathbf{H}(T)\mathbf{u}(0)\\
\mathbf{x}(2T) &= \mathbf{G}(T)\mathbf{x}(T) + \mathbf{H}(T)\mathbf{u}(T)\\
&\cdots\\
\mathbf{x}(nT) &= \mathbf{G}(T)\mathbf{x}((n-1)T) + \mathbf{H}(T)\mathbf{u}((n-1)T) = \mathbf{0}
\end{aligned}$$

Para ser uma solução prática, o vetor de controle $\mathbf{u}(kT)$ deve ser determinado como uma função do vetor de estado $\mathbf{x}(kT)$.

Note que uma vez que supusemos a norma de $\mathbf{u}(kT)$ não-limitada, podemos obter o tempo de resposta menor fazendo o tempo de amostragem mais curto. Note também que se o sinal de controle é limitado, então o número de períodos de amostragem necessário para transferir o estado inicial arbitrário para a origem pode aumentar.

Ilustraremos a determinação do $\mathbf{u}(kT)$ requerido usando um exemplo.

Exemplo 16.9 Considere o seguinte sistema de tempo discreto:

$$\begin{bmatrix} x_1((k+1)T) \\ x_2((k+1)T) \end{bmatrix} = \begin{bmatrix} 1 & 1-e^{-T} \\ 0 & e^{-T} \end{bmatrix} \begin{bmatrix} x_1(kT) \\ x_2(kT) \end{bmatrix} + \begin{bmatrix} e^{-T}+T-1 \\ 1-e^{-T} \end{bmatrix} [u(kT)] \tag{16.23}$$

Note que esta é a versão discreta do sistema de tempo contínuo descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [u]$$

Vamos reescrever a Eq. (16.23) como

$$\mathbf{x}((k+1)T) = \mathbf{G}(T)\mathbf{x}(kT) + \mathbf{H}(T)u(kT)$$

Supomos que $u(kT)$ é não-limitada, ou

$$-\infty \leq u(kT) \leq \infty$$

Naturalmente, este sistema é de estado completamente controlável. Resolveremos o problema de transferir um estado inicial arbitrário para a origem usando um sinal de controle constante por trechos.

Como o sistema dado é de segunda-ordem, necessitamos no máximo de dois períodos de amostragem para transferir qualquer estado inicial $\mathbf{x}(0)$ para a origem. Observando que

$$\begin{aligned}
\mathbf{x}(T) &= \mathbf{G}(T)\mathbf{x}(0) + \mathbf{H}(T)u(0)\\
\mathbf{x}(2T) &= \mathbf{0} = \mathbf{G}(T)\mathbf{x}(T) + \mathbf{H}(T)u(T)\\
&= \mathbf{G}^2(T)\mathbf{x}(0) + \mathbf{G}(T)\mathbf{H}(T)u(0) + \mathbf{H}(T)u(T)
\end{aligned} \tag{16.24}$$

obtemos

$$\mathbf{x}(0) = -\mathbf{G}^{-1}(T)\mathbf{H}(T)u(0) - \mathbf{G}^{-2}(T)\mathbf{H}(T)u(T) \tag{16.25}$$

Substituindo a Eq. (16.25) na Eq. (16.24), obtemos

$$\mathbf{x}(T) = -\mathbf{G}^{-1}(T)\mathbf{H}(T)u(T) \tag{16.26}$$

Observando que

$$\mathbf{G}^{-1}(T)\mathbf{H}(T) = \begin{bmatrix} 1 & 1-e^{T} \\ 0 & e^{T} \end{bmatrix} \begin{bmatrix} e^{-T}+T-1 \\ 1-e^{-T} \end{bmatrix} = \begin{bmatrix} T-e^{T}+1 \\ e^{T}-1 \end{bmatrix}$$

$$\mathbf{G}^{-2}(T)\mathbf{H}(T) = \begin{bmatrix} 1 & 1-e^{2T} \\ 0 & e^{2T} \end{bmatrix} \begin{bmatrix} e^{-T}+T-1 \\ 1-e^{-T} \end{bmatrix} = \begin{bmatrix} T+e^{T}-e^{2T} \\ e^{2T}-e^{T} \end{bmatrix}$$

obtemos das Eqs. (16.25) e (16.26)

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = -\begin{bmatrix} T-e^{T}+1 \\ e^{T}-1 \end{bmatrix}[u(0)] - \begin{bmatrix} T+e^{T}-e^{2T} \\ e^{2T}-e^{T} \end{bmatrix} u(T)] \tag{16.27}$$

$$\begin{bmatrix} x_1(T) \\ x_2(T) \end{bmatrix} = -\begin{bmatrix} T-e^{T}+1 \\ e^{T}-1 \end{bmatrix}[u(T)] \tag{16.28}$$

Combinando as Eqs. (16.27) e (16.28), temos

$$\begin{bmatrix} x_1(0) & x_1(T) \\ x_2(0) & x_2(T) \end{bmatrix} = \begin{bmatrix} -T+e^{T}-1 & -T-e^{T}+e^{2T} \\ -e^{T}+1 & -e^{2T}+e^{T} \end{bmatrix} \begin{bmatrix} u(0) & u(T) \\ u(T) & 0 \end{bmatrix}$$

Portanto

$$\begin{bmatrix} u(0) & u(T) \\ u(T) & 0 \end{bmatrix} = \begin{bmatrix} -T+e^{T}-1 & -T-e^{T}+e^{2T} \\ -e^{T}+1 & -e^{2T}+e^{T} \end{bmatrix}^{-1} \begin{bmatrix} x_1(0) & x_1(T) \\ x_2(0) & x_2(T) \end{bmatrix} \tag{16.29}$$

A Eq. (16.29) define a seqüência de controle ótimo $u(0)$ e $u(T)$. Note que $u(T)$ pode ser expresso em termos de $\mathbf{x}(0)$ apenas ou $\mathbf{x}(T)$ apenas.

Para ilustrar a solução $u(kT)$ em termos de $\mathbf{x}(kT)$, vamos supor que T é igual à 1 segundo. Então,

$$\begin{bmatrix} u(0) & u(1) \\ u(1) & u(0) \end{bmatrix} = \begin{bmatrix} -1{,}58 & -1{,}24 \\ 0{,}58 & 0{,}24 \end{bmatrix} \begin{bmatrix} x_1(0) & x_1(1) \\ x_2(0) & x_2(1) \end{bmatrix} \tag{16.30}$$

do que obtemos

$$u(k) = -1{,}58x_1(k) - 1{,}24x_2(k) \qquad (k = 0, 1)$$

Com esta lei de controle, qualquer estado inicial $\mathbf{x}(0)$ pode ser transferido para a origem em 2 segundos. A Fig. 16.1 é o diagrama de blocos do sistema de controle ótimo.

Vamos agora determinar os estados iniciais a partir dos quais podemos transferir o sistema para a origem em um período de amostragem. Para tal, fazemos $\mathbf{x}(T) = \mathbf{0}$ na Eq. (16.24). Então,

$$\mathbf{x}(T) = \mathbf{G}(T)\mathbf{x}(0) + \mathbf{H}(T)u(0) = \mathbf{0}$$

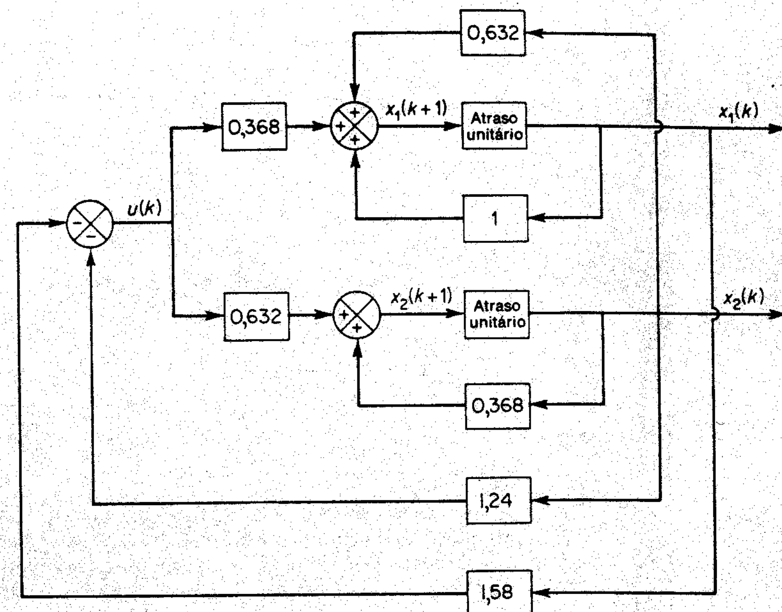


Fig. 16.1 Diagrama de blocos do sistema de controle ótimo considerado no Exemplo 16.9.

do qual obtemos

$$\mathbf{x}(0) = -\mathbf{G}^{-1}(T)\mathbf{H}(T)u(0) = -\begin{bmatrix} T-e^{T}+1 \\ e^{T}-1 \end{bmatrix} u(0)$$

Para $T = 1$,

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 0{,}72 \\ -1{,}72 \end{bmatrix}[u(0)] \tag{16.31}$$

Da Eq. (16.31), achamos que se o estado inicial está na reta

$$1{,}72x_1(0) + 0{,}72x_2(0) = 0$$

então ele pode ser transferido para a origem em um período de amostragem, ou 1 segundo. Alternativamente impomos que o estado inicial seja transferido para a origem em dois períodos de amostragem, ou 2 segundos.

Finalmente, note que se queremos reduzir o tempo de resposta para no máximo 0,5 segundo, isto pode ser feito escolhendo-se $T = 0{,}25$ segundo. Entretanto, isto aumentará a amplitude do sinal de controle.

Controle de tempo ótimo de sistemas de tempo contínuo com sinais de controle limitados. Usando argumentos plausíveis e experimentação, acreditava-se no passado que se um sistema de controle está sendo operado sob potência limitada, então

podia ser levado de um estado para outro no mínimo tempo utilizando-se sempre toda a potência disponível. Esta hipótese é o chamado princípio "bangue-bangue". Embora muitos sistemas de controle de tempo ótimo com potência de controle limitada sejam do tipo bangue-bangue, não é verdade que todos os sistemas são deste tipo. (Por exemplo, se a função de transferência de um processo inclui zeros, pode-se mostrar que a solução estrita de bangue-bangue em que o sinal de controle chaveia entre valores constantes $+u_0$ e $-u_0$ não é ótimo.)

Considere um processo descrito por

$$\overset{(n)}{y} + a_1 \overset{(n-1)}{y} + \cdots + a_{n-1}\dot{y} + a_n y = bu \tag{16.32}$$

ou

$$\frac{Y(s)}{U(s)} = \frac{b}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}$$

onde a amplitude de u é limitada, ou

$$-u_0 \leq u \leq u_0$$

Pode-se provar (Referências B-2 e L-4) que, para o controle de tempo ótimo do processo dado pela Eq. (16.32), a lei de controle é do tipo bangue-bangue; isto é, o sinal de controle toma o valor máximo, cujo sinal é chaveado no máximo $n - 1$ vezes.

Na análise que se segue, limitaremos nossa discussão a processos descritos pelas Eq. (16.32). Não provaremos o princípio do bangue-bangue aqui. Em vez disso, usaremos o fato já provado de que o controle de tempo ótimo de sistemas descritos pela Eq. (16.32) é do tipo bangue-bangue e utilizaremos isto na solução do problema específico onde $n = 2$.

Para sistemas de segunda-ordem, a abordagem por plano de fase é o mais conveniente para a determinação da curva de chaveamento em que o sinal do sinal de controle deve ser modificado. Para sistemas de ordem superior, entretanto, qualquer que seja a abordagem, é bastante difícil determinar as superfícies de chaveamento no espaço de estados n-dimensional. Na maioria dos casos, a construção explícita das superfícies de chaveamento no espaço de estados n-dimensional é praticamente impossível.

Exemplo 16.10 Considere o controle de tempo ótimo do processo descrito por $1/Js^2$. O controlador ótimo é do tipo bangue-bangue ou de máximo esforço. Isto é, o controlador age no motor para fornecer o máximo torque positivo para acelerar a carga ou o máximo torque negativo para desacelerar a carga, dependendo de o erro ser positivo ou negativo. Supondo que o sistema está sujeito apenas a mudanças nas condições iniciais, determine a curva de chaveamento no plano de fase.

Referindo-se à Fig. 16.2, vemos que a equação para o sistema é

$$J\ddot{c} = u = \pm T \tag{16.33}$$

Como o sistema está sujeito apenas a mudanças nas condições iniciais, fazemos $r = 0$ para $t > 0$. Então obtemos $c = -e$, e a Eq. (16.33) se torna

$$\ddot{e} = -\frac{u}{J} = \mp\frac{T}{J}$$

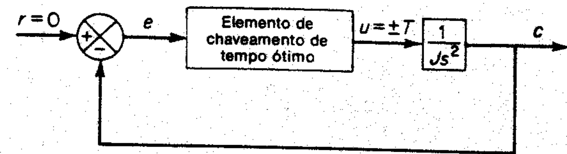


Fig. 16.2 Diagrama de blocos de um sistema de controle de tempo ótimo.

Esta última equação pode ser reescrita como

$$\dot{e}\frac{d\dot{e}}{de} = \mp\frac{T}{J}$$

da qual obtemos

$$\dot{e}^2 = \mp 2\frac{T}{J}e + K$$

onde K é uma constante. Esta equação descreve duas famílias de parábolas no plano $e - \dot{e}$, dependendo de $u = T$ ou $u = -T$.

Uma vez que colocamos $u = T$ ou $u = -T$, a parábola particular é determinada pelo valor de K, que depende das condições iniciais $e(0)$ e $\dot{e}(0)$. As únicas parábolas que passam pela origem são as parábolas AOA' e a parábola BOB', vistas na Fig. 16.3.

Referindo-se à Fig. 16.3 vemos que, para alcançar a origem, um ponto representativo do plano e—$\dot{e}$ deve seguir a trajetória AO ou BO. Isto significa que, dado o estado inicial representado pelo ponto P na Fig. 16.4, devemos usar $u = T$ até que o ponto representativo, seguindo a trajetória parabólica, alcança o ponto Q, onde o sinal de controle deve ser chaveado para $u = -T$. Então o ponto representativo se move ao longo da trajetória QO. Isto dará o menor tempo de resposta, ou a resposta de tempo ótimo.

Para ver isto graficamente, refira-se à Fig. 16.5. Vamos comparar as duas trajetórias PQO e $PQRSTUO$. Notando que o tempo requerido para ir do ponto Q ao ponto O é o mesmo para ir do ponto S ao ponto T, vemos que

$$(t_f - t_i)_{PQO} < (t_f - t_i)_{PQRSTUO}$$

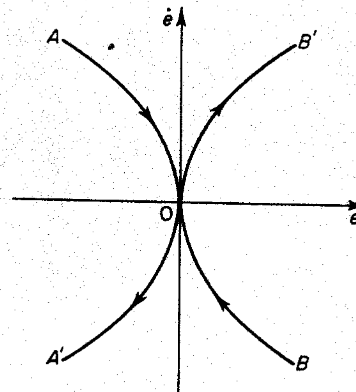


Fig. 16.3 Diagrama de plano de fase mostrando trajetórias passando através da origem.

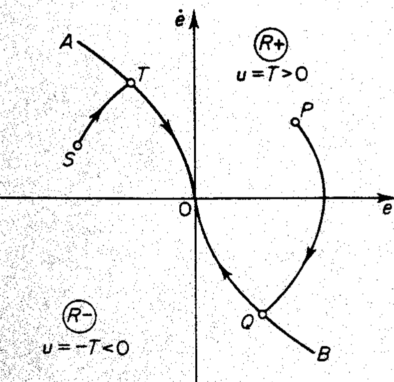


Fig. 16.4 Diagrama de plano de fase mostrando trajetórias de tempo ótimo.

onde $t_f - t_i$ é o tempo necessário para ir do ponto P à origem.

De forma semelhante, comparando as duas trajetórias PQO e $PRSTO$ na Fig. 16.6, vemos que leva mais tempo para se ir do ponto R ao ponto T através do ponto S do que através do ponto Q, porque a velocidade média é maior para trajetória RQT do que para a trajetória RST. Portanto,

$$(t_f - t_i)_{PQO} < (t_f - t_i)_{PRSTO}$$

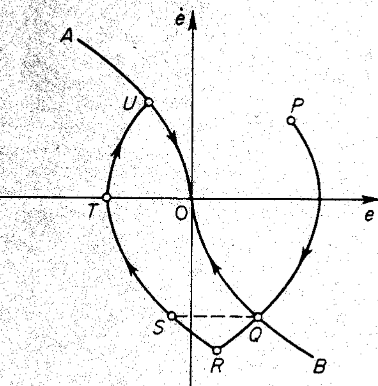


Fig. 16.5 Diagrama de plano de fase mostrando trajetórias ótimas e não-ótimas.

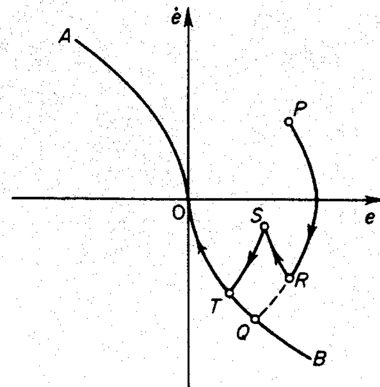


Fig. 16.6 Diagrama de plano de fase mostrando trajetórias ótimas e não-ótimas.

Portanto, se o chaveamento ocorre em pontos outros que os da parábola AOB, a trajetória requer um tempo extra para se alcançar a origem e não pode ser de tempo ótimo.

De forma semelhante, se o estado inicial é o ponto S na Fig. 16.4, então a trajetória ótima é STO. Portanto, para a resposta de tempo ótimo, um ponto representativo deve permanecer na sua parábola inicial até que alcance a parábola AO ou a parábola BO. O sinal de controle, então, tem seu sinal chaveado, e o ponto representativo segue a trajetória ao longo da parábola AO ou parábola BO até que a origem é alcançada.

Resumindo, o controle de tempo ótimo do processo cuja função de transferência é dada por $1/Js^2$ foi analisado. Determinamos que a curva de chaveamento separa o plano e–$\dot{e}$ em regiões R_+ e R_-, como visto na Fig. 16.4, onde R_+ é o conjunto de pontos para a direita e R_- é o conjunto de pontos para a esquerda da curva de chaveamento AOB. Uma formalização da lei de controle para este problema pode ser dada como segue: A lei de controle de tempo ótimo em função do estado é dada por

$u(t) = T$ para (x_1, x_2) na região R_+ ou sobre a curva de chaveamento AO
$u(t) = -T$ *para* (x_1, x_2) na região R_- ou sobre a curva de chaveamento BO

É importante observar que, independente do estado inicial, impomos no máximo um chaveamento do sinal de controle.

Finalmente, do ponto de vista prático, notamos o seguinte:

1. Efeitos como atrito podem modificar a curva de chaveamento, e chaveamentos imprecisos podem levar a uma trajetória que não se aproxima da origem do plano de fase. Por esta razão, uma região de operação linear pode ser provida nas vizinhanças da origem. Um sistema como este é freqüentemente chamado de um sistema de controle de tempo ótimo de modo dual.
2. Para simplificar o problema de implementação do controlador ótimo, podemos aproximar a curva de chaveamento parabólica por uma série de segmentos de reta. O sistema resultante não é de tempo ótimo, mas estará perto do ótimo. (Veja Problema A.16.5.)

Alguns comentários sobre o controle de tempo ótimo de sistemas de ordem elevada com sinais de controle limitados. Considere o seguinte sistema de ordem elevada com sinal de controle limitado:

$$\overset{(n)}{y} + a_1 \overset{(n-1)}{y} + \cdots + a_{n-1}\dot{y} + a_n y = bu$$

ou na representação por espaço de estados

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ b \end{bmatrix} [u]$$

$$y = [1 \quad 0 \quad \cdots \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

O problema de controle de tempo ótimo deste sistema é o de transferir o estado a partir de um certo ponto do espaço de estados para a origem (ou algum outro ponto especificado) no tempo mínimo que é consistente com as dadas restrições. Como as condições iniciais podem ser arbitrárias, qualquer ponto no espaço de estados pode ser um ponto inicial. Portanto, é necessário determinar a lei de controle que minimiza o tempo de resposta como uma função das variáveis de estado. Isto requer a determinação do lugar geométrico dos pontos no espaço de estados em que o sinal de controle deve ser chaveado para a direção oposta.

O uso de um método matricial fornecerá a solução do problema de chaveamento de tais sistemas de ordem elevada em termos gerais. Deve-se mencionar que, na prática, o problema real não é o de achar a solução em forma matricial geral, mas encontrar um método para resolver várias equações simultâneas envolvendo termos transcendentais em um tempo razoável. Quando se trabalha com sistemas de ordem elevada, notamos que as computações envolvidas se tornam bastante trabalhosas à medida que o número de variáveis ou a ordem do sistema aumenta. (O método matricial não oferece nenhum auxílio para resolver este problema real.)

16.5 SISTEMAS DE CONTROLE ÓTIMO BASEADOS NOS ÍNDICES DE DESEMPENHO QUADRÁTICO

Nesta seção consideraremos o projeto de sistemas de controle ótimo baseados nos índices de desempenho quadráticos. O sistema de controle que vamos considerar pode ser representado por

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu}$$

onde

$\mathbf{x}$ = vetor de estado (vetor real n-dimensional)
$\mathbf{u}$ = vetor de controle (vetor real r-dimensional)
$\mathbf{A}$ = matriz constante $n \times n$
$\mathbf{B}$ = matriz constante $n \times r$

Ao projetar sistemas de controle, estamos muitas vezes interessados em escolher o vetor de controle $\mathbf{u}(t)$ de tal forma que um dado índice de desempenho é minimizado. Pode-se provar que os índices de desempenho quadráticos onde os limites de integração são 0 e ∞, tais como

$$J = \int_0^\infty L(\mathbf{x}, \mathbf{u})dt$$

onde $L(\mathbf{x}, \mathbf{u})$ é uma função quadrática de $\mathbf{x}$ e $\mathbf{u}$, fornecerão leis de controle lineares, isto é,

$$\mathbf{u}(t) = -\mathbf{Kx}(t)$$

onde $\mathbf{K}$ é uma matriz $r \times n$, ou

$$\begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_r \end{bmatrix} = - \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1n} \\ k_{21} & k_{22} & \cdots & k_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ k_{r1} & k_{r2} & \cdots & k_{rn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

Portanto, o projeto de sistemas de controle ótimo baseados em tais índices de desempenho quadráticos se resume na determinação dos elementos da matriz $\mathbf{K}$.

A seguir, discutiremos inicialmente o problema de otimização de parâmetros, isto é, a determinação dos valores ótimos dos parâmetros do sistema. Especificamente, consideramos um sistema com estado inicialmente fora da origem

$$\dot{\mathbf{x}} = \mathbf{Ax}, \qquad \mathbf{x}(0) = \mathbf{c}$$

onde $\mathbf{A}$ tem um ou mais parâmetros ajustáveis. Deseja-se transferir qualquer estado inicial para a origem enquanto minimizamos o índice de desempenho quadrático

$$J = \int_0^\infty \mathbf{x}'\mathbf{Qx}\, dt$$

onde $\mathbf{Q}$ é uma matriz positiva-definida (ou positiva-semidefinida) real e simétrica. O problema então se torna o de determinar o(s) valor(es) do(s) parâmetro(s) ajustável(eis) de tal forma a minimizar o índice de desempenho.

A seguir consideraremos o problema de controle ótimo baseado em índices de desempenho quadráticos e determinaremos a lei de controle. Isto é, consideramos o problema de determinar o vetor de controle ótimo $\mathbf{u}(t)$ para o sistema descrito por

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu}$$

e o índice de desempenho dado por

$$J = \int_0^\infty (\mathbf{x}'\mathbf{Qx} + \mathbf{u}'\mathbf{Ru})\, dt$$

onde $\mathbf{Q}$ é uma matriz positiva-definida (ou positiva-semidefinida) real e simétrica, $\mathbf{R}$ é uma matriz positiva-definida real e simétrica, e $\mathbf{u}$ não sofre restrições.

Há várias abordagens diferentes para a solução destes dois tipos de problemas. Nesta seção, apresentamos uma baseada no segundo método de Liapunov. (Lembrar que no Cap. 7 discutimos uma abordagem clássica para a solução de problemas semelhantes de otimização.)

Otimização de sistemas de controle através do segundo método de Liapunov. Classicamente, sistemas de controle são inicialmente projetados e então sua estabilidade é examinada. Uma abordagem diferente desta é onde as condições para estabilidade são formuladas em primeiro lugar e então o sistema é projetado dentro destas limitações. Para uma grande classe de sistemas de controle, pode-se mostrar uma relação direta entre funções de Liapunov e os índices de desempenho quadráticos generalizados usados na síntese de sistemas de controle ótimo. Se o segundo

método de Liapunov é utilizado para formar a base do projeto de um controlador ótimo, então estamos seguros de que o sistema funcionará, isto é, que a saída do sistema será continuamente levada em direção do seu valor desejado. Portanto o sistema projetado tem uma configuração com características inerentes de estabilidade.

Problemas de otimização de parâmetros resolvidos usando-se o segundo método de Liapunov. Em seguida, discutiremos uma relação direta entre funções de Liapunov e índices de desempenho quadráticos generalizados e resolveremos o problema de otimização de parâmetros usando esta relação. Vamos considerar o sistema:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$$

onde todos os autovalores de **A** têm partes reais negativas, ou a origem $\mathbf{x} = \mathbf{0}$ é assintoticamente estável. (Chamaremos uma matriz como esta **A** de uma matriz *estável*.) Queremos minimizar o índice de desempenho quadrático definido por

$$J = \int_0^\infty \mathbf{x}'\mathbf{Q}\mathbf{x}\, dt$$

onde **Q** é uma matriz positiva-definida (ou positiva-semidefinida) real e simétrica.

Mostraremos agora que uma função de Liapunov pode realmente ser usada na solução deste problema. Vamos supor que

$$\mathbf{x}'\mathbf{Q}\mathbf{x} = -\frac{d}{dt}(\mathbf{x}'\mathbf{P}\mathbf{x})$$

onde **P** é uma matriz positiva-definida real e simétrica. Então, obtemos

$$\mathbf{x}'\mathbf{Q}\mathbf{x} = -\dot{\mathbf{x}}'\mathbf{P}\mathbf{x} - \mathbf{x}'\mathbf{P}\dot{\mathbf{x}} = -\mathbf{x}'\mathbf{A}'\mathbf{P}\mathbf{x} - \mathbf{x}'\mathbf{P}\mathbf{A}\mathbf{x} = -\mathbf{x}'(\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A})\mathbf{x}$$

Pelo segundo método de Liapunov, sabemos que para um dado **Q** há uma matriz **P**, se **A** for estável, tal que

$$\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q} \qquad (16.34)$$

Portanto podemos determinar os elementos de **P** desta equação.

O índice de desempenho J pode ser calculado como

$$J = \int_0^\infty \mathbf{x}'\mathbf{Q}\mathbf{x}\, dt = -\mathbf{x}'\mathbf{P}\mathbf{x}\Big|_0^\infty = -\mathbf{x}'(\infty)\mathbf{P}\mathbf{x}(\infty) + \mathbf{x}'(0)\mathbf{P}\mathbf{x}(0)$$

Como todos os autovalores de **A** têm partes reais negativas, temos $\mathbf{x}(\infty) \to \mathbf{0}$. Portanto, obtemos

$$J = \mathbf{x}'(0)\mathbf{P}\mathbf{x}(0)$$

Portanto o índice de desempenho J pode ser obtido em termos da condição inicial $\mathbf{x}(0)$ e de **P**, que está relacionado com **A** e **Q** pela Eq. (16.34). Se, por exemplo, um parâmetro do sistema deve ser ajustado para minimizar o índice de desempenho J, então isto pode ser conseguido minimizando-se $\mathbf{x}'(0)\mathbf{P}\mathbf{x}(0)$ com relação ao parâmetro em questão. Como $\mathbf{x}(0)$ é a dada condição inicial e **Q** também é dada, **P** é uma função dos elementos de **A**. Portanto este processo de minimização resultará no

valor ótimo do parâmetro ajustável.

É importante notar que o valor ótimo deste parâmetro depende, em geral, da condição inicial $\mathbf{x}(0)$. Entretanto, se $\mathbf{x}(0)$ envolve apenas um componente não nulo, isto é, $x_1(0) \neq 0$, e outras condições iniciais são nulas, então o valor ótimo do parâmetro não depende do valor numérico de $x_1(0)$. (Veja o exemplo seguinte.)

Exemplo 16.11 Considere o sistema visto na Fig. 16.7. Determine o valor do coeficiente de amortecimento $\zeta > 0$ de tal forma que, quando o sistema é sujeito a uma entrada degrau unitário $r(t) = 1(t)$, o seguinte índice de desempenho é minimizado:

$$J = \int_0^\infty \mathbf{x}'(t)\mathbf{Q}\mathbf{x}(t)\, dt$$

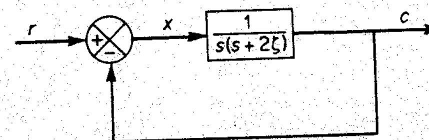


Fig. 16.7 Sistema de controle.

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x \\ \dot{x} \end{bmatrix}, \qquad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix} \qquad (\mu > 0)$$

Supõe-se que o sistema está inicialmente em repouso.

Da Fig. 16.7, obtemos a seguinte equação para o sistema:

$$\ddot{c} + 2\zeta\dot{c} + c = r$$

Notando que $x = r - c$, $r(t) = 1(t)$, e que as condições iniciais são nulas, temos

$$\ddot{x} + 2\zeta\dot{x} + x = 0 \qquad (t > 0)$$

A representação de espaço de estados desta última equação se torna

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -2\zeta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

ou

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$$

onde

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & -2\zeta \end{bmatrix}$$

Como **A** é uma matriz estável, o valor de J é dado por

$$J = \mathbf{x}'(0)\mathbf{P}\mathbf{x}(0)$$

onde **P** é determinado de

$$\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}$$

ou

$$\begin{bmatrix} 0 & -1 \\ 1 & -2\zeta \end{bmatrix}\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & -2\zeta \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -\mu \end{bmatrix}$$

Esta equação resulta nas seguintes três equações:

$$-2p_{12} = -1$$
$$p_{11} - 2\zeta p_{12} - p_{22} = 0$$
$$2p_{12} - 4\zeta p_{22} = -\mu$$

Resolvendo estas três equações para obter os p_{ij}, obtemos

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} \zeta + \dfrac{1+\mu}{4\zeta} & \dfrac{1}{2} \\ \dfrac{1}{2} & \dfrac{1+\mu}{4\zeta} \end{bmatrix}$$

Portanto o índice de desempenho J se torna

$$J = \mathbf{x}'(0)\mathbf{P}\mathbf{x}(0)$$
$$= \left(\zeta + \frac{1+\mu}{4\zeta}\right)x_1^2(0) + x_1(0)x_2(0) + \frac{1+\mu}{4\zeta}x_2^2(0)$$

Substituindo as condições iniciais $x_1(0) = 1$, $x_2(0) = 0$ nesta última equação, obtemos

$$J = \zeta + \frac{1+\mu}{4\zeta}$$

Para minimizar J com relação a ζ, fazemos $\partial J/\partial\zeta = 0$, ou

$$\frac{\partial J}{\partial \zeta} = 1 - \frac{1+\mu}{4\zeta^2} = 0$$

Isto nos dá

$$\zeta = \frac{\sqrt{1+\mu}}{2}.$$

Portanto o valor ótimo de ζ é $\sqrt{1+\mu}/2$. Por exemplo, se $\mu = 1$, então o valor ótimo de ζ é $\sqrt{2}/2$ ou 0,707.

Sistemas de controle ótimo baseados em índices de desempenho quadráticos. Dada a equação do sistema

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{16.35}$$

consideraremos o problema de controle ótimo de determinar a matriz **K** do vetor de controle ótimo

$$\mathbf{u}(t) = -\mathbf{K}\mathbf{x}(t) \tag{16.36}$$

de tal modo a minimizar o índice de desempenho

$$J = \int_0^\infty (\mathbf{x}'\mathbf{Q}\mathbf{x} + \mathbf{u}'\mathbf{R}\mathbf{u})\,dt \tag{16.37}$$

onde **Q** é uma matriz positiva-definida (ou positiva-semidefinida) real e simétrica e **R** uma matriz positiva-definida real e simétrica. Note que o segundo termo do lado direito da Eq. (16.37) se relaciona com o gasto de energia dos sinais de controle. As matrizes **Q** e **R** determinam a importância relativa do erro e o gasto desta energia. Neste problema, supomos que o vetor de controle $\mathbf{u}(t)$ não sofre restrições.

Como mencionado anteriormente, sem demonstração, a lei de controle dada pela Eq. (16.36) é a lei de controle ótimo. Portanto, se os elementos desconhecidos da matriz **K** são determinados de tal forma a minimizar o índice de desempenho, então $\mathbf{u}(t) = -\mathbf{K}\mathbf{x}(t)$ é ótimo para qualquer estado inicial $\mathbf{x}(0)$. O diagrama de bloco mostrando a configuração ótima é visto na Fig. 16.8.

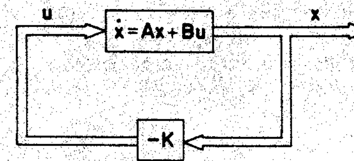


Fig. 16.8 Sistema de controle ótimo.

Vamos agora resolver o problema de otimização. Substituindo a Eq. (16.36) na Eq. (16.35), obtemos

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} - \mathbf{B}\mathbf{K}\mathbf{x} = (\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{x}$$

Nas derivações seguintes, suporemos que a matriz $\mathbf{A} - \mathbf{B}\mathbf{K}$ é estável, ou que os autovalores de $\mathbf{A} - \mathbf{B}\mathbf{K}$ têm partes reais negativas.

Substituindo a Eq. (16.36) na Eq. (16.37), obtemos

$$J = \int_0^\infty (\mathbf{x}'\mathbf{Q}\mathbf{x} + \mathbf{x}'\mathbf{K}'\mathbf{R}\mathbf{K}\mathbf{x})\,dt$$
$$= \int_0^\infty \mathbf{x}'(\mathbf{Q} + \mathbf{K}'\mathbf{R}\mathbf{K})\mathbf{x}\,dt$$

Seguindo a discussão dada para resolver o problema de otimização de parâmetros, fazemos

$$\mathbf{x}'(\mathbf{Q} + \mathbf{K}'\mathbf{R}\mathbf{K})\mathbf{x} = -\frac{d}{dt}(\mathbf{x}'\mathbf{P}\mathbf{x})$$

Então, obtemos

$$\mathbf{x}'(\mathbf{Q} + \mathbf{K}'\mathbf{R}\mathbf{K})\mathbf{x} = -\dot{\mathbf{x}}'\mathbf{P}\mathbf{x} - \mathbf{x}'\mathbf{P}\dot{\mathbf{x}}$$
$$= -\mathbf{x}'[(\mathbf{A} - \mathbf{B}\mathbf{K})'\mathbf{P} + \mathbf{P}(\mathbf{A} - \mathbf{B}\mathbf{K})]\mathbf{x}$$

Comparando ambos os lados desta última equação e notando que esta equação deve ser verdadeira para qualquer **x**, impomos que

$$(\mathbf{A} - \mathbf{B}\mathbf{K})'\mathbf{P} + \mathbf{P}(\mathbf{A} - \mathbf{B}\mathbf{K}) = -(\mathbf{Q} + \mathbf{K}'\mathbf{R}\mathbf{K}) \tag{16.38}$$

Pelo segundo método de Liapunov, se $\mathbf{A} - \mathbf{BK}$ é uma matriz estável, então existe uma matriz positiva definida **P** que satisfaz a Eq. (16.38). Então, notando que $\mathbf{x}(\infty) = \mathbf{0}$, o índice de desempenho pode ser escrito como

$$J = \mathbf{x}'(0)\mathbf{P}\mathbf{x}(0) \tag{16.39}$$

As etapas de projeto podem ser escritas como:

1. Determinar a matriz **P** que satisfaz a Eq. (16.38) como uma função de **K**.
2. Substituir a matriz **P** na Eq. (16.39). Então o índice de desempenho se torna uma função de **K**.
3. Determinar os elementos da matriz **K** de tal forma que o índice de desempenho J seja minimizado. A minimização de J em relação aos elementos k_{ij} de **K** pode ser obtida fazendo-se $\partial J/\partial k_{ij}$ igual a zero e resolvendo as equações, determinando os valores ótimos de k_{ij}.

Para detalhes deste método de projeto, veja o Problema A.16.8. Quando o número de elementos k_{ij} não é pequeno, este método não é conveniente.

Dispõe-se de um método melhor para o projeto do sistema de controle ótimo. Como se supôs que **R** é uma matriz positiva definida real e simétrica, podemos escrever

$$\mathbf{R} = \mathbf{T}'\mathbf{T}$$

onde **T** é uma matriz não-singular. Então a Eq. (16.38) pode ser escrita como

$$(\mathbf{A}' - \mathbf{K}'\mathbf{B}')\mathbf{P} + \mathbf{P}(\mathbf{A} - \mathbf{B}\mathbf{K}) + \mathbf{Q} + \mathbf{K}'\mathbf{T}'\mathbf{T}\mathbf{K} = \mathbf{0}$$

que pode ser reescrita como

$$\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A} + [\mathbf{T}\mathbf{K} - (\mathbf{T}')^{-1}\mathbf{B}'\mathbf{P}]'[\mathbf{T}\mathbf{K} - (\mathbf{T}')^{-1}\mathbf{B}'\mathbf{P}] - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{P} + \mathbf{Q} = \mathbf{0}$$

A minimização de J com relação a **K** requer a minimização de

$$\mathbf{x}'[\mathbf{T}\mathbf{K} - (\mathbf{T}')^{-1}\mathbf{B}'\mathbf{P}]'[\mathbf{T}\mathbf{K} - (\mathbf{T}')^{-1}\mathbf{B}'\mathbf{P}]\mathbf{x}$$

com relação a **K**. Como este valor é não-negativo, o mínimo ocorre quando este se anula, ou quando

$$\mathbf{T}\mathbf{K} = (\mathbf{T}')^{-1}\mathbf{B}'\mathbf{P}$$

Portanto

$$\mathbf{K} = \mathbf{T}^{-1}(\mathbf{T}')^{-1}\mathbf{B}'\mathbf{P} = \mathbf{R}^{-1}\mathbf{B}'\mathbf{P} \tag{16.40}$$

A Eq. (16.40) nos dá a matriz ótima **K**. A matriz **P** na Eq. (16.40) deve satisfazer a Eq. (16.38) ou a seguinte equação reduzida:

$$\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}'\mathbf{P} + \mathbf{Q} = \mathbf{0} \tag{16.41}$$

A Eq. (16.41) é chamada a equação de Riccati de matriz reduzida. As etapas de projeto podem ser escritas como:

1. Resolver a Eq. (16.41), a equação de Riccati de matriz reduzida, para a matriz **P**.
2. Substituir esta matriz **P** na Eq. (16.40). A matriz resultante **K** é a ótima.

Um exemplo de projeto baseado nesta última abordagem é dado no Exemplo 16.12.

Se a matriz $\mathbf{A} - \mathbf{BK}$ é estável, o método presente sempre nos dá o resultado correto. Kalman mostrou que a imposição de a matriz $\mathbf{A} - \mathbf{BK}$ ser estável equivale à de que a característica da matriz

$$[\mathbf{S}' \mid \mathbf{A}'\mathbf{S}' \mid \cdots \mid (\mathbf{A}')^{n-1}\mathbf{S}'] \tag{16.42}$$

seja igual a n, onde a matriz **S** é definida por

$$\mathbf{S}'\mathbf{S} = \mathbf{Q}$$

Esta condição na característica pode ser aplicada convenientemente para testar se a matriz $\mathbf{A} - \mathbf{BK}$ é estável.

Entretanto é importante notar que, mesmo que $\mathbf{A} - \mathbf{BK}$ não seja uma matriz estável, o método presente ainda nos dará resultados corretos em um caso especial em que a matriz **P** determinada da Eq. (16.41) se torna positiva semidefinida; não obstante ela nos fornece $\mathbf{x}'(\infty)\mathbf{P}\mathbf{x}(\infty) = 0$.

Finalmente, note que se o índice de desempenho é dado em termos do vetor de saída ao invés do vetor de estado, isto é,

$$J = \int_0^\infty (\mathbf{y}'\mathbf{Q}\mathbf{y} + \mathbf{u}'\mathbf{R}\mathbf{u})\, dt$$

então o índice pode ser modificado usando-se a equação de saída

$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

para

$$J = \int_0^\infty (\mathbf{x}'\mathbf{C}'\mathbf{Q}\mathbf{C}\mathbf{x} + \mathbf{u}'\mathbf{R}\mathbf{u})\, dt$$

e o mesmo procedimento que discutimos aqui pode ser aplicado para se obter a matriz ótima **K**.

Exemplo 16.12 Considere o sistema visto na Fig. 16.9. Supondo que o sinal de controle seja

$$u(t) = -\mathbf{K}\mathbf{x}(t)$$

projete a matriz de ganho de realimentação ótima **K** de tal forma que o seguinte índice de desempenho seja minimizado:

$$J = \int_0^\infty (\mathbf{x}'\mathbf{Q}\mathbf{x} + u'u)\, dt$$

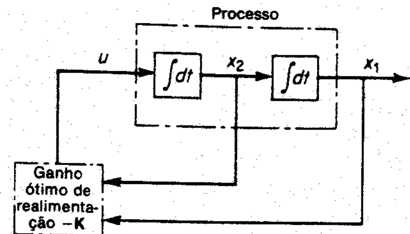


Fig. 16.9 Sistema de controle.

onde

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix} \qquad (\mu \geq 0)$$

Da Fig. 16.9, vemos que a equação para o processo é

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [u]$$

Observando que

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\mu} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\mu} \end{bmatrix} = \mathbf{S'S}$$

determinamos a característica da matriz dada por (16.42), ou

$$[\mathbf{S}' \mid \mathbf{A'S'}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{\mu} & 1 & 0 \end{bmatrix}$$

igual a dois. Portanto $\mathbf{A} - \mathbf{BK}$ é uma matriz estável. e a abordagem por Liapunov apresentada nesta seção fornece o resultado correto.

Demonstraremos o uso da equação de Riccati de matriz reduzida no projeto do sistema de controle ótimo. Vamos resolver a Eq. (16.41), reescrita como

$$\mathbf{A'P} + \mathbf{PA} - \mathbf{PBR}^{-1}\mathbf{B'P} + \mathbf{Q} = \mathbf{0}$$

ou

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$
$$- \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} [1] [0 \quad 1] \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Esta equação pode ser simplificada para

$$\begin{bmatrix} 0 & 0 \\ p_{11} & p_{12} \end{bmatrix} + \begin{bmatrix} 0 & p_{11} \\ 0 & p_{12} \end{bmatrix} - \begin{bmatrix} p_{12}^2 & p_{12}p_{22} \\ p_{12}p_{22} & p_{22}^2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

de que obtemos as seguintes três equações:

$$1 - p_{12}^2 = 0$$
$$p_{11} - p_{12}p_{22} = 0$$
$$\mu + 2p_{12} - p_{22}^2 = 0$$

Resolvendo estas três equações simultâneas para p_{11}, p_{12} e p_{22}, impondo que **P** seja positiva definida, obtemos

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} \sqrt{\mu + 2} & 1 \\ 1 & \sqrt{\mu + 2} \end{bmatrix}$$

A matriz de ganho de realimentação ótima **K** é então obtida de

$$\begin{aligned} \mathbf{K} &= \mathbf{R}^{-1}\mathbf{B'P} \\ &= [1] [0 \quad 1] \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \\ &= [p_{12} \quad p_{22}] \\ &= [1 \quad \sqrt{\mu + 2}] \end{aligned}$$

Portanto o sinal de controle ótimo é

$$u = -\mathbf{Kx} = -x_1 - \sqrt{\mu + 2}x_2 \tag{16.43}$$

Note que a lei de controle dada pela Eq. (16.43) fornece um resultado ótimo para qualquer estado inicial sob o dado índice de desempenho. A Fig. 16.10 é o diagrama de bloco para este sistema.

Comentários conclusivos

1. A característica de uma lei de controle ótimo baseada em um índice de desempenho quadrático é que ela é uma função linear das variáveis de estado, que implica que necessitamos realimentar todas as variáveis de estado. Isto requer que todas estas variáveis estejam disponíveis para a realimentação. Portanto é desejável representar o sistema em termos de variáveis de estado

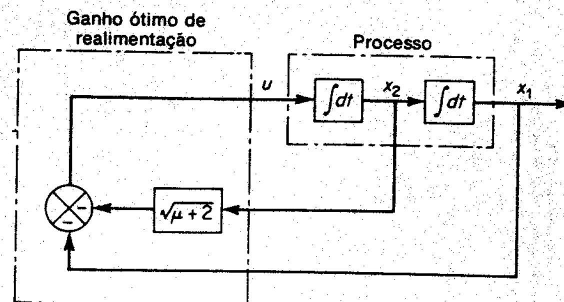


Fig. 16.10 Controle ótimo do processo visto na Fig. 16.9.

mensuráveis. (Em sistemas complicados, é pouco provável que possamos medir todas as variáveis de estado. Então devemos estimar as não mensuráveis e usar estes valores estimados para gerar sinais de controle ótimo.) É importante notar que devemos evitar diferenciar uma variável de estado para gerar outra. A diferenciação de um sinal sempre diminui a relação sinal-ruído porque o ruído em geral tem flutuações mais rápidas do que o sinal de comando. Às vezes a relação sinal-ruído pode ser diminuída substancialmente usando-se apenas um processo de diferenciação. Geralmente é desejável evitar processos de diferenciação; por exemplo, se a aceleração do deslocamento de saída é desejada, é melhor medir a aceleração diretamente usando um acelerômetro ao invés de diferenciar o sinal de velocidade.

2. Quando o sistema de controle ótimo é projetado no domínio do tempo, é desejável investigar as características de resposta em freqüência para compensar efeitos de ruído. As características de resposta em freqüência do sistema devem ser tais que o sistema atenue bastante na gama de freqüências em que ruído e ressonância de componentes são esperados. (Para compensar efeitos de ruído devemos, em alguns casos, ou modificar a configuração ótima e aceitar um desempenho subótimo ou modificar o índice de desempenho.)

3. Se o limite superior de integração do índice de desempenho J dado pela Eq. (16.37) é finito, então pode-se mostrar que o vetor de controle ótimo ainda é uma função linear das variáveis de estado e as entradas de referência, mas com coeficientes variantes no tempo. (Portanto, a determinação do vetor de controle ótimo envolve a de matrizes variantes no tempo ótimas.)

16.6 SISTEMAS DE CONTROLE DE REFERÊNCIA-MODELO

No Cap. 10, apresentamos técnicas de projeto e compensação de sistemas de controle lineares e invariantes no tempo. Como todos os processos físicos têm certo grau de não-linearidade, o sistema projetado se comportará satisfatoriamente apenas em uma gama limitada de operação.

Se a suposição de linearidade na equação do processo é eliminada, as técnicas de projeto do Cap. 10 não são aplicáveis. Em tal caso o método de referência-modelo para o projeto de sistemas apresentado nesta seção pode ser útil.

Sistemas de controle de referência-modelo. Um dos métodos úteis para especificar o desempenho de um sistema é através de um modelo que produzirá a saída desejada para uma dada entrada. O modelo não necessita ser em *hardware*. Ele pode ser apenas um modelo matemático simulado em um computador. Em um sistema de controle de referência-modelo, a saída do modelo e a do processo são comparadas e a diferença é utilizada para gerar os sinais de controle.

O controle por referência-modelo tem sido usado para se obter desempenhos aceitáveis em algumas situações muito difíceis de controle envolvendo não-linearidades e/ou parâmetros variantes no tempo.

Projeto de um controlador.* Suporemos que o processo é caracterizado pela seguinte equação de estado:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{16.44}$$

* Referências M-8 e V-1.

onde

$\mathbf{x}$ = vetor de estado (vetor real n-dimensional)
$\mathbf{u}$ = vetor de controle (vetor real r-dimensional)
$\mathbf{f}$ = função vetorial

Deseja-se que o sistema de controle siga de perto algum sistema-modelo. Nosso problema de projeto é o de sintetizar um controlador que sempre gera um sinal que força o estado do processo sempre em direção ao estado do modelo. A Fig. 16.11 é o diagrama de blocos mostrando a configuração do sistema.

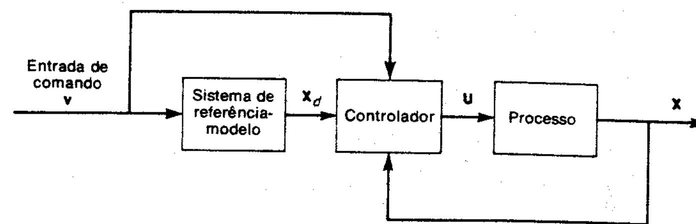


Fig. 16.11 Sistema de controle de referência-modelo.

Suporemos que o sistema de referência-modelo é linear e descrito por

$$\dot{\mathbf{x}}_d = \mathbf{A}\mathbf{x}_d + \mathbf{B}\mathbf{v} \tag{16.45}$$

onde

$\mathbf{x}_d$ = vetor de estado do modelo (vetor real n-dimensional)
$\mathbf{v}$ = vetor de entrada (vetor real r-dimensional)
$\mathbf{A}$ = matriz constante $n \times n$
$\mathbf{B}$ = matriz constante $n \times r$

Supomos que os autovalores de $\mathbf{A}$ têm partes reais negativas, de tal forma que o sistema referência-modelo tem um estado de equilíbrio assintoticamente estável.

Vamos definir o vetor de erro $\mathbf{e}$ por

$$\mathbf{e} = \mathbf{x}_d - \mathbf{x} \tag{16.46}$$

No problema presente, queremos reduzir o vetor de erro para zero usando um vetor de controle adequado $\mathbf{u}$. Das Eqs. (16.44), (16.45) e (16.46), obtemos

$$\begin{aligned}\dot{\mathbf{e}} = \dot{\mathbf{x}}_d - \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x}_d + \mathbf{B}\mathbf{v} - \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \\ &= \mathbf{A}\mathbf{e} + \mathbf{A}\mathbf{x} - \mathbf{f}(\mathbf{x}, \mathbf{u}, t) + \mathbf{B}\mathbf{v}\end{aligned} \tag{16.47}$$

A Eq. (16.47) é uma equação diferencial para o vetor de erro.

Agora projetaremos um controlador tal que no estado estacionário $\mathbf{x} = \mathbf{x}_d$ e $\dot{\mathbf{x}} = \dot{\mathbf{x}}_d$, ou $\mathbf{e} = \dot{\mathbf{e}} = 0$. Portanto a origem $\mathbf{e} = 0$ será um estado de equilíbrio.

Um ponto de partida conveniente na síntese do vetor de controle $\mathbf{u}$ é a

construção de uma função de Liapunov para o sistema dada pela Eq. (16.47).

Vamos supor que a forma da função de Liapunov é

$$V(\mathbf{e}) = \mathbf{e}'\mathbf{P}\mathbf{e}$$

onde **P** é uma matriz positiva-definida real e simétrica. Tomando a derivada de $V(\mathbf{e})$ com relação ao tempo, obtemos

$$\begin{aligned}\dot{V}(\mathbf{e}) &= \dot{\mathbf{e}}'\mathbf{P}\mathbf{e} + \mathbf{e}'\mathbf{P}\dot{\mathbf{e}} \\ &= [\mathbf{e}'\mathbf{A}' + \mathbf{x}'\mathbf{A}' - \mathbf{f}'(\mathbf{x}, \mathbf{u}, t) + \mathbf{v}'\mathbf{B}']\mathbf{P}\mathbf{e} - \mathbf{e}'\mathbf{P}[\mathbf{A}\mathbf{e} + \mathbf{A}\mathbf{x} - \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \\ &\quad + \mathbf{B}\mathbf{v}] = \mathbf{e}'(\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A})\mathbf{e} + 2M \qquad (16.48)\end{aligned}$$

onde

$$M = \mathbf{e}'\mathbf{P}[\mathbf{A}\mathbf{x} - \mathbf{f}(\mathbf{x}, \mathbf{u}, t) + \mathbf{B}\mathbf{v}] = \text{grandeza escalar}$$

A função suposta $V(\mathbf{e})$ é uma função de Liapunov se

1. $\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}$ é uma matriz negativa-definida.
2. O vetor de controle **u** pode ser escolhido para tornar a grandeza escalar M não positiva.

Então, notando que $V(\mathbf{e}) \to \infty$ quando $\| \mathbf{e} \| \to \infty$, vemos que o estado de equilíbrio $\mathbf{e} = \mathbf{0}$ é assintoticamente estável globalmente. A Condição 1 pode sempre ser satisfeita através de uma escolha adequada de **P**, uma vez que se supõe os autovalores de **A** têm partes reais negativas. O problema aqui é o de escolher um vetor de controle **u** tal que M seja zero ou negativo.

Ilustraremos a aplicação do método presente para o projeto de um controlador não-linear usando um exemplo.

Exemplo 16.13 Considere um processo não-linear variante no tempo descrito por

$$\begin{bmatrix}\dot{x}_1 \\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}0 & 1 \\ -b & -a(t)x_2\end{bmatrix}\begin{bmatrix}x_1 \\ x_2\end{bmatrix} + \begin{bmatrix}0 \\ 1\end{bmatrix}[u]$$

onde $a(t)$ é variante no tempo e b uma constante positiva. Supondo que a equação da referência-modelo é

$$\begin{bmatrix}\dot{x}_{d1} \\ \dot{x}_{d2}\end{bmatrix} = \begin{bmatrix}0 & 1 \\ -\omega_n^2 & -2\zeta\omega_n\end{bmatrix}\begin{bmatrix}x_{d1} \\ x_{d2}\end{bmatrix} + \begin{bmatrix}0 \\ \omega_n^2\end{bmatrix}[v] \qquad (16.49)$$

projete um controlador não-linear que resultará em uma operação estável do sistema.

Defina o vetor de erro por

$$\mathbf{e} = \mathbf{x}_d - \mathbf{x}$$

e uma função de Liapunov por

$$V(\mathbf{e}) = \mathbf{e}'\mathbf{P}\mathbf{e}$$

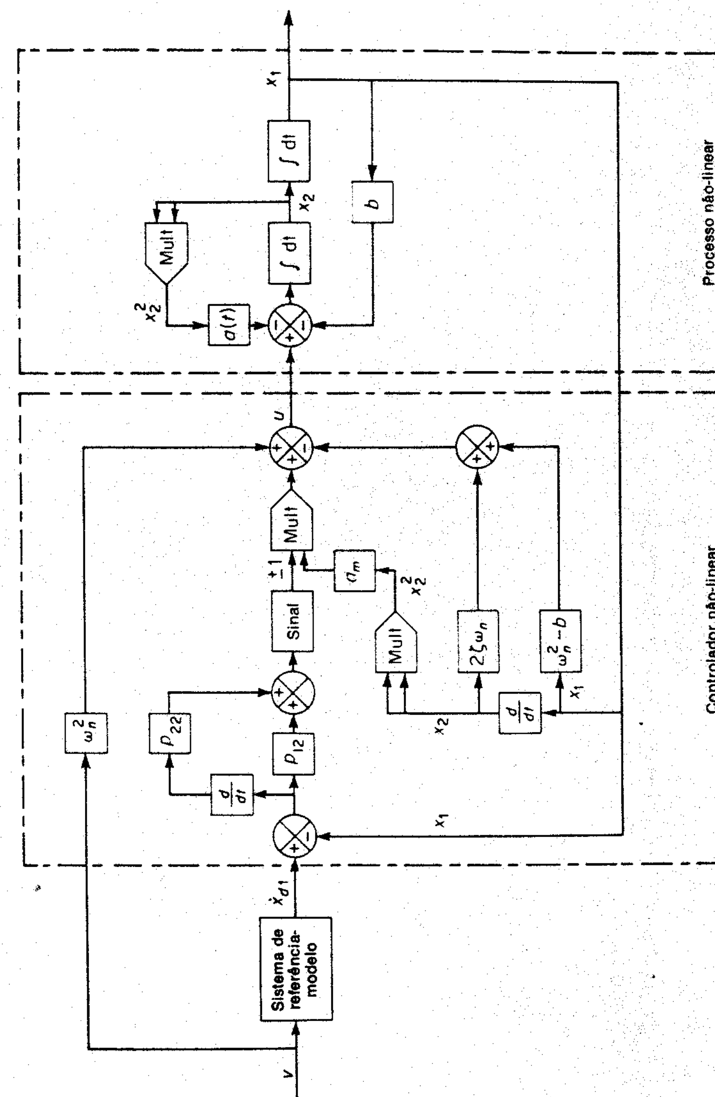


Fig. 16.12 Controle de referência-modelo de um processo não-linear.

onde **P** é uma matriz positiva-definida real e simétrica. Então, referindo-se à Eq. (16.48), obtemos $\dot{V}(\mathbf{e})$ como

$$\dot{V}(\mathbf{e}) = \mathbf{e}'(\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A})\mathbf{e} + 2M$$

onde

$$M = \mathbf{e}'\mathbf{P}[\mathbf{A}\mathbf{x} - \mathbf{f}(\mathbf{x}, \mathbf{u}, t) + \mathbf{B}\mathbf{v}]$$

Identificando as matrizes **A** e **B** da Eq. (16.49) e escolhendo a matriz **Q** como

$$\mathbf{Q} = \begin{bmatrix} q_{11} & 0 \\ 0 & q_{22} \end{bmatrix} = \text{positiva definida}$$

obtemos

$$\dot{V}(\mathbf{e}) = -(q_{11}e_1^2 + q_{22}e_2^2) + 2M$$

onde

$$\begin{aligned} M &= [e_1 \quad e_2]\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}\left\{\begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\zeta\omega_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right. \\ &\quad \left. - \begin{bmatrix} 0 & 1 \\ -b & -a(t)x_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0 \\ u \end{bmatrix} + \begin{bmatrix} 0 \\ \omega_n^2 v \end{bmatrix}\right\} \\ &= (e_1p_{12} + e_2p_{22})[-(\omega_n^2 - b)x_1 - 2\zeta\omega_n x_2 + a(t)x_2^2 + \omega_n^2 v - u] \end{aligned}$$

Se escolhermos u tal que

$$u = -(\omega_n^2 - b)x_1 - 2\zeta\omega_n x_2 + \omega_n^2 v + a_m x_2^2 \,\text{sinal}(e_1p_{12} + e_2p_{22}) \qquad (16.50)$$

onde

$$a_m = \text{máx.}|a(t)|$$

então

$$M = (e_1p_{12} + e_2p_{22})[a(t) - a_m \,\text{sinal}(e_1p_{12} + e_2p_{22})]x_2^2 = \text{não positiva}$$

Com a função de controle u dada pela Eq. (16.50), o estado de equilíbrio $\mathbf{e} = \mathbf{0}$ é assintoticamente estável globalmente. Portanto a Eq. (16.50) define uma lei de controle não linear que fornecerá uma operação assintoticamente estável. O diagrama em blocos para o sistema de controle presente é visto na Fig. 16.12.

Note que a taxa de convergência da resposta transitória depende da matriz **P**, que por sua vez depende da matriz **Q** escolhida no início do projeto.

16.7 SISTEMAS DE CONTROLE ADAPTATIVO

Em anos recentes, o interesse em sistemas de controle adaptativos aumentou rapidamente, juntamente com o interesse e progresso em tópicos de controle em geral. O termo *sistema adaptativo* tem uma variedade de significados específicos,

mas normalmente implica que o sistema é capaz de se acomodar a mudanças imprevisíveis no meio-ambiente, venham estas mudanças de dentro do sistema ou de fora. Este conceito é de muito interesse para o projetista de sistema, pois, um sistema altamente adaptativo, além de se acomodar para mudanças ambientais, também se acomodaria para erros de projeto de engenharia e compensaria por falhas ou incertezas moderadas de componentes secundários do sistema, aumentando portanto a confiabilidade do sistema.

Inicialmente apresentaremos alguns conceitos básicos de sistemas de controle adaptativo e explicaremos o que tais sistemas representam. Então discutiremos as funções necessárias que um controlador deve desempenhar para que seja chamado adaptativo. Finalmente, introduziremos alguns conceitos de sistemas com aprendizado.

Introdução. Na maioria de sistemas de controle e realimentação, pequenos desvios nos valores de parâmetros dos seus valores projetados não causarão quaisquer problemas na operação normal do sistema, desde que estes parâmetros estejam dentro da malha. Entretanto, se os parâmetros do processo variam muito com mudanças ambientais, então o sistema de controle pode apresentar uma resposta satisfatória para uma condição ambiental mas deixar de ter um desempenho satisfatório sob outras condições. Em certos casos, grandes variações de parâmetros de processo podem até causar instabilidade.

Na análise mais simples, podem-se considerar diferentes conjuntos de valores dos parâmetros do processo. É então desejável projetar um sistema de controle que funciona bem para todos os conjuntos. Tão logo esta imposição é formulada, o problema de controle ótimo estrito perde sua importância. Impondo um bom desempenho em uma certa gama, abandonamos o melhor desempenho para um conjunto de parâmetros.

Se a função de transferência do processo pode ser identificada continuamente, então podemos compensar variações na função de transferência do processo simplesmente variando parâmetros ajustáveis do controlador, e desta forma obter um desempenho satisfatório do sistema continuamente sob várias condições ambientais. Tal abordagem adaptativa é bastante útil para lidar com um problema onde o processo é normalmente exposto a ambientes variáveis de tal forma que parâmetros do processo mudam de tempo em tempo. (Como não são previsíveis mudanças na maioria dos casos práticos, um controlador de parâmetro fixo ou um controlador variante no tempo pré-programado não pode resolver o problema.)

Definição de sistemas de controle adaptativo. A adaptação é uma característica fundamental de organismos vivos pois eles tentam manter o equilíbrio fisiológico diante de condições ambientais variantes. Uma abordagem para o projeto de sistemas adaptativos é então considerar os aspectos adaptativos do comportamento humano ou animal e desenvolver sistemas que se comportam aproximadamente de forma análoga.

Há diferentes definições de sistemas de controle adaptativo atualmente em uso na literatura. O caráter um tanto vago da maioria das definições e classificações de sistemas adaptativos é devido à grande variedade de mecanismos através dos quais a adaptação pode ser conseguida, e também devido a uma falha em diferenciar entre as manifestações externas do comportamento adaptativo e os mecanismos internos usados para consegui-lo. As diferentes definições acontecem primordialmente por causa das várias classificações e delineações que dividem os sistemas de controle em adaptativos e não-adaptativos. (O pequeno grau de adaptabilidade necessária pela maioria das especificações de sistemas poderia ser obtido usando-se técnicas familiares de realimentação com ganhos fixos, compensadores e, em alguns casos,

não linearidades.)

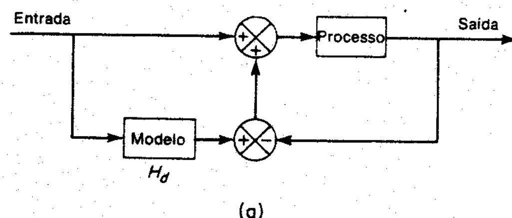
Acharemos necessário definir características de sistemas adaptativos que são fundamentalmente diferentes daquelas de sistemas convencionais de realimentação, de tal forma que podemos restringir nossa atenção apenas àqueles aspectos específicos do comportamento e projeto de sistemas adaptativos. Neste livro definimos sistemas de controle adaptativos como segue:

Definição. Um sistema de controle adaptativo é aquele que mede, de forma contínua e automática, as características dinâmicas (tal como a função de transferência) do processo, as compara com as características dinâmicas desejadas, e usa a diferença para variar parâmetros ajustáveis do sistema (normalmente características do controlador) ou para gerar um sinal atuante de tal forma que o desempenho ótimo pode ser mantido independentemente das mudanças ambientais; alternativamente, tal sistema pode continuamente medir seu próprio desempenho de acordo com um dado índice de desempenho e modificar, se necessário, seus próprios parâmetros, de tal forma a manter desempenho ótimo independentemente de mudanças ambientais.

Para usarmos o termo sistema adaptativo, devem existir características de auto-organização. Se o ajuste dos parâmetros do sistema é feito apenas por medição direta do ambiente, o sistema não é adaptativo.

Um sistema aparentemente adaptativo é o exemplo do autopiloto de um avião, que é projetado para ajustar seus ganhos de malha em função da altitude a fim de compensar as correspondentes mudanças nos parâmetros do avião. O ajuste é baseado diretamente na informação sobre o ambiente (neste caso, a pressão atmosférica) e não em um esquema de auto-organização. Estes sistemas não possuem quaisquer características de auto-organização, e portanto são essencialmente sistemas convencionais a malha-fechada.

Um outro exemplo de sistemas que podem parecer adaptativos, mas na realidade não o são, se encontra no campo dos sistemas de referência-modelo. Alguns destes sistemas (tal como o considerado no Exemplo 16.13) meramente usam a diferença entre a resposta do modelo e a resposta do processo como um sinal de entrada para o processo, como visto na Fig. 16.13(a). Estes sistemas não podem ser



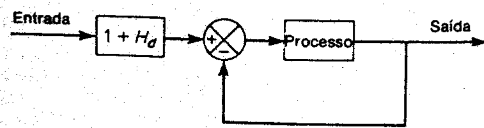


Fig. 16.13 (a) Sistema de referência-modelo; (b) diagrama de bloco simplificado.

considerados como verdadeiramente adaptativos uma vez que, através de manipulação de diagramas de blocos, reduzimos esta configuração para aquela da Fig. 16.13(b), que é simplesmente a de uma malha de realimentação básica com um filtro antes da malha. (Note que alguns autores chamam este tipo de controle de *adaptação a modelo*. O modelo ou é um modelo físico ou um sistema simulado no computador. O modelo não tem parâmetros variáveis.)

Índices de desempenho. A base do controle adaptativo está na premissa de que há alguma condição de operação ou desempenho para o sistema melhor que qualquer outra. Portanto se torna necessário definir o que constitui um desempenho ótimo. Em sistemas de controle adaptativo, tal desempenho é definido em termos do índice de desempenho, que devemos escolher após definir nossos objetivos. Estes objetivos são tão diversos como os sistemas para os quais são aplicados, mas geralmente é possível generalizar o objetivo da otimização como o de minimização do custo de operação, ou a maximização do lucro.

Algumas características gerais geralmente consideradas desejáveis são

1. confiabilidade
2. seletividade
3. aplicabilidade

Portanto, o índice de desempenho deve ser confiável, ou deve ser uma medida uniforme de "qualidade" para sistemas de todas as ordens. Ele deve ter seletividade, ou envolver um ótimo definido com precisão em função dos parâmetros do sistema. Não deve ter ótimos locais ou pontos de sela. O índice de desempenho deve ser facilmente aplicável para sistemas práticos e facilmente mensurável.

Se o índice de desempenho toma o valor zero na condição de operação ótima, ao invés de um máximo ou mínimo, então pode ser usado como o sinal de erro da malha adaptativa e para realimentação diretamente em alguns sistemas.

Note que, em geral, todos os índices de desempenho matematicamente tratáveis têm uma séria desvantagem em comum; embora especifiquem o custo de operação do sistema em termos de erro e energia, não nos dão informação sobre as características de resposta transitória do sistema. Portanto, um sistema que é projetado para operar de forma ótima do ponto de vista de máximo "lucro" pode ter características transitórias indesejáveis, ou mesmo ser instável. Desta forma, para assegurar características de resposta satisfatórias, podemos precisar de critérios secundários relacionados a características de resposta para poder influenciar na escolha dos elementos de ponderação do custo.

Finalmente, deve-se lembrar que o índice de desempenho usado em um sistema de controle adaptativo define um desempenho ótimo para aquele sistema. Isto significa que o índice de desempenho essencialmente nos dá o limite superior do desempenho do sistema. Portanto, a seleção de um índice de desempenho adequado é de fundamental importância.

Controladores adaptativos. Um controlador adaptativo pode consistir nas seguintes três funções:

1. Identificação das características dinâmicas do processo.
2. Ato de decisão baseado na identificação do processo.
3. Modificação ou atuação baseada na decisão tomada.

Se o processo é conhecido apenas imperfeitamente, talvez por causa de parâmetros aleatórios variando no tempo ou em virtude dos efeitos de mudanças ambientais nas características dinâmicas do processo, então a identificação inicial, a decisão, e os procedimentos de modificação não serão suficientes para minimizar (ou maximizar) o índice de desempenho. Então se torna necessário realizar estes procedimentos continuamente ou em intervalos de tempo, dependendo de quão

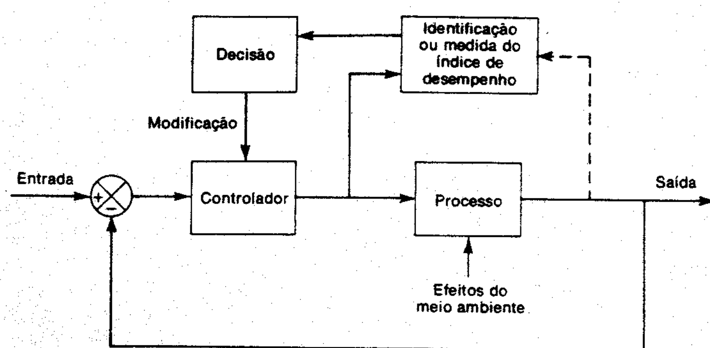


Fig. 16.14 Representação em diagrama em blocos de um sistema de controle adaptativo.

rapidamente os parâmetros do processo estão mudando. Este constante "reprojetamento" ou auto-organização do sistema para compensar as mudanças imprevisíveis no processo é o aspecto do desempenho normalmente considerado ao se definir um sistema de controle adaptativo.

Uma representação em diagramas em bloco de um sistema de controle adaptativo é vista na Fig. 16.14. Neste sistema, o processo é identificado e o índice de desempenho medido contínua ou periodicamente. Uma vez que isto foi feito, o índice de desempenho é comparado com o ótimo e uma decisão é tomada com base nos achados de como modificar o sinal atuante. Como o processo é identificado dentro do sistema em si, o ajuste dos parâmetros é uma operação a malha-fechada. Note que usando esta adaptação a malha-fechada, o problema da estabilidade pode surgir.

A seguir, explicaremos em algum detalhe as três funções: identificação, decisão, e modificação.

Identificação das características dinâmicas do processo. As características dinâmicas do processo devem ser medidas e identificadas continuamente ou, pelo menos, muito freqüentemente. Isto deve ser conseguido sem afetar a operação normal do sistema. Para identificar as características de um sistema, devemos fazer um teste e analisar os resultados. (Para um sistema de controle, isto implica a imposição de um sinal de controle no processo e a análise da resposta do sistema.) A identificação pode ser feita a partir de dados de operação normal do processo ou pelo uso de sinais de teste, tais como os sinais senoidais de baixa amplitude ou de vários sinais estocásticos de pequena amplitude. Na prática, a aplicação direta de entradas em degrau ou impulsivas não pode ser feita. (Exceto em certos casos especiais, o processo estará em operação normal durante o teste de tal forma que os sinais de teste impostos não devem perturbar as saídas normais; além do mais, entradas normais e ruído no sistema não devem perturbar e confundir o teste.) As entradas normais são ideais como sinais de teste uma vez que não encontramos as dificuldades de saídas não desejadas ou entradas perturbadas. Entretanto, a identificação com entradas normais só é possível quando elas têm características de sinal convenientes (banda, amplitude etc.) para a identificação adequada.

Os sinais de teste estocásticos são bastante convenientes em certas aplicações. Usando técnicas de correlação cruzada, podemos analisar a saída como uma função da entrada estocástica para determinar as características de resposta. Com uma

entrada aleatória, a energia de excitação se espalha sobre uma banda de freqüências fazendo o efeito tolerável. Além do mais, o dispositivo de cálculo da correlação cruzada pode ser projetado para manter o nível do sinal de teste baixo.

A identificação não pode demorar muito pois, caso contrário, podem ocorrer novas variações dos parâmetros do processo. O tempo de identificação deve ser suficientemente curto comparado com a taxa de mudanças ambientais. Com o tempo de identificação limitado, geralmente é impossível identificar o processo completamente; o melhor que se pode esperar é apenas uma identificação parcial.

E importante notar que nem todos os sistemas adaptativos requerem uma identificação de forma explícita. Alguns sistemas já foram identificados a ponto de a medida do valor do índice de desempenho poder indicar que parâmetros do controlador devem ser modificados. Isto é, o sistema é muito bem conhecido, de tal forma que uma medida do índice de desempenho completa a identificação.

Por outro lado, se a identificação do processo é muito difícil, devemos medir diretamente o índice de desempenho e construir um controlador adaptativo baseado nele. Se não se necessita a identificação e a adaptabilidade é baseada apenas em medidas do índice de desempenho, o sistema de controle é chamado de um sistema de controle otimalizante. Como é obtida uma auto-organização usando-se este método, consideraremos tais sistemas como adaptativos.

A dificuldade de fazer uma identificação realística dependerá de quanta informação requeremos sobre o processo e da quantidade de conhecimento prévio deste. Em geral, estes são também os fatores que determinarão se devemos usar uma abordagem de identificação ou uma busca direta no espaço de parâmetros do controlador em função do índice de desempenho, que é discutido mais tarde sob o título de sistemas de controle otimalizantes.

Decisão baseada na identificação do processo. Decisão se refere aqui a uma decisão baseada nas características do processo que foram identificadas e no valor calculado do índice de desempenho.

Uma vez que o processo foi identificado, ele é comparado com as características ótimas (ou desempenho ótimo), e então uma decisão deve ser tomada de como os parâmetros ajustáveis (características do controlador) devem ser variados para manter um desempenho ótimo. A decisão é tomada por um computador.

Modificação baseada na decisão tomada. Modificação se refere à modificação de sinais de controle de acordo com os resultados da identificação e da decisão. Na maioria dos casos, a decisão e a modificação são conceitualmente uma única operação, com a modificação consistindo em um meio de mecanizar a transformação de um sinal de saída de decisão em um sinal de controle (a entrada do processo).

Este sinal de controle, o sinal de entrada para o processo, pode ser modificado de duas formas. O primeiro método é o de ajustar os parâmetros do controlador para compensar as mudanças na dinâmica do processo. Isto é chamado de modificação de parâmetro do controlador. O segundo método é o de sintetizar o sinal de controle ótimo, baseado na função de transferência do processo, no índice de desempenho e na resposta transitória desejada. Isto é chamado de síntese de sinal de controle.

A escolha entre a modificação de parâmetros do controlador e a síntese de sinal de controle é primordialmente de *hardware*, uma vez que os dois métodos são conceitualmente equivalentes. Quando a confiabilidade é muito importante, como em aplicações aeroespaciais, o uso da adaptação por mudança de parâmetros é freqüentemente favorecido ao invés da síntese de sinal de controle. (Isto acontece porque o sistema pode operar mesmo após a falha da malha adaptativa se o sinal de controle não é inteiramente dependente da porção adaptativa do sistema.) Aperfei-

çoamentos em componentes, especialmente em equipamentos eletrônicos complexos, podem fazer a preferência passar para a síntese de sinais.

Sistemas de controle otimalizantes. Os sistemas de controle otimalizantes se apóiam bastante em técnicas de otimização. De forma geral, a otimização consiste na busca no espaço de parâmetros variáveis do controlador em função de algum índice de desempenho para determinar onde este é maximizado ou minimizado. Está implícito na afirmação anterior o fato de que um índice de desempenho escalar, que é uma função das saídas do sistema, pode ser definido de tal forma que seu extremo representa o melhor desempenho possível do sistema. Isto é geralmente possível e necessário para qualquer sistema de controle adaptativo.

Os métodos de achar o ponto de operação ótimo são basicamente procedimentos de tentativa e erro. No método de "descida mais abrupta" (*steepest-descent*), o gradiente da superfície do índice de desempenho é medido observando-se os efeitos de pequenas mudanças nos parâmetros variáveis. (Isto pode ser chamado de método de sentir a derivada.) O vetor de parâmetros é então movido em direção à máxima inclinação, ou de uma quantidade fixa, ou de uma quantidade determinada pelo gradiente da superfície. Nos locais onde os parâmetros variam lentamente, o gradiente pode ser calculado relativamente com pouca freqüência. Em condições de variações mais rápidas dos parâmetros, entretanto, um procedimento conhecido como polarização alternada é superior ao método de sentir a derivada. Com polarização alternada, o sistema nunca é operado na condição ótima, mas operado alternadamente a uma distância fixa em ambos os lados do ótimo calculado, e um novo ótimo é calculado a partir da diferença nos valores do índice de desempenho.

Provavelmente a maior vantagem do método do controle otimalizante é que não se colocam restrições no processo. O processo pode ser não-linear, de múltipla entrada e múltipla saída, variante no tempo etc. Uma grande dificuldade deste método de otimização é que não se determinou nenhum método satisfatório para discriminar entre extremos locais e extremos globais. Portanto, esta abordagem é útil para qualquer processo físico cuja superfície de desempenho tem um único ótimo e cujas variações são lentas o suficiente para que o sistema de controle se acomode a elas.

Sistemas com aprendizado. Uma diferença significativa entre um operador humano treinado e um controlador adaptativo discutido acima é que o operador humano reconhece entradas familiares e pode usar suas experiências passadas aprendidas para reagir de uma forma ótima. Sistemas de controle adaptativo são projetados para modificar o sinal de controle à medida que o meio externo ao sistema muda de tal forma a manter um desempenho ótimo.

Um sistema que é capaz de reconhecer características e padrões familiares de uma situação e que usa suas experiências passadas aprendidas para se comportar de uma forma ótima é chamado de um *sistema com aprendizado*.

Um sistema com aprendizado é um sistema de nível superior ao dos sistemas adaptativos. Gibson divide o espaço de todos os sistemas de controle em quatro níveis básicos de hierarquia:

1. malha aberta
2. malha fechada
3. malha adaptativa
4. malha de aprendizado

onde cada nível é sensível a um índice de desempenho ou erro de controle medido no próximo nível mais baixo e onde existirão níveis mais altos do que o quarto para ambientes mais complexos.

Um sistema com aprendizado respondendo a uma situação familiar não vai requerer a identificação do sistema. A abordagem para o projeto de tal sistema é a de "ensinar" ao sistema qual a melhor escolha para cada situação. Uma vez que o sistema aprendeu a lei de controle ótimo para cada situação possível, ele pode operar perto da condição ótima independentemente de mudanças ambientais.

Um sistema com aprendizado, quando sujeito a uma nova situação, aprende como se comportar através de um método adaptativo. Se o sistema é sujeito a uma mesma situação que aprendeu antes, ele reconhecerá isto e se comportará de forma ótima sem ter que passar pelo mesmo método adaptativo. Modelos de comportamento humano que estão sendo desenvolvidos por muitos pesquisadores proverão, sem dúvida, resultados úteis para aplicações em sistemas com aprendizado.

Comentários conclusivos. Desenvolvimentos recentes de veículos aeroespaciais de grande desempenho e processos de alta eficiência impõem restrições cada vez maiores nos seus sistemas de controle associados. Ao projetar tais sistemas de controle, nos preocupamos com o projeto de sistemas que satisfarão as especificações impostas pelos usuários sob as condições antecipadas de operação. A maioria dos sistemas de controle que exigem grande desempenho para uma grande gama de condições de operação, serão necessariamente adaptativos até um certo grau. Quando um alto grau de adaptabilidade é claramente requerido, a maioria das especificações atuais serão obedecidas por um sistema do tipo identificação-decisão-modificação com modificação seqüencial ou contínua, dependendo da taxa de variação dos parâmetros variantes.

Os desenvolvimentos mais fascinantes em sistemas de controle adaptativo estão na área de reconhecimento de padrões e de sistemas com aprendizado. As técnicas de reconhecimento de padrões poderão algum dia responder à necessidade de uma técnica geral de identificação. Quando acopladas com abordagens de aprendizado, eles podem reduzir consideravelmente a "barreira de tempo" que no momento atrapalha muitos sistemas de controle adaptativos.

Esta seção apresentou apenas um esboço de sistemas de controle adaptativo. O leitor interessado em tais sistemas deve-se referir a resultados recentes de pesquisa disponível na literatura.

PROBLEMAS ILUSTRATIVOS E SOLUÇÕES

Problema A.16.1 Mostre que o seguinte sistema:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u$$
$$y = \mathbf{C}\mathbf{x}$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} -1 & -2 & -2 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{C} = [1 \quad 1 \quad 0]$$

é de estado completamente controlável e completamente observável.

Solução. Para controlabilidade completa de estado, os vetores **B**, **AB** e $\mathbf{A}^2\mathbf{B}$ devem ser linearmente independentes. Para este sistema, obtemos

$$\mathbf{B} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{AB} = \begin{bmatrix} -4 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{A}^2\mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ -5 \end{bmatrix}$$

Como

$$\begin{vmatrix} 2 & -4 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & -5 \end{vmatrix} = -10$$

a característica da matriz $[\mathbf{B}|\mathbf{AB}|\mathbf{A}^2\mathbf{B}]$ é 3, e o sistema é de estado completamente controlável.

Para observabilidade completa, os vetores $\mathbf{C}'$, $\mathbf{A}'\mathbf{C}'$, e $(\mathbf{A}')^2\mathbf{C}'$ devem ser linearmente independentes. Para este sistema,

$$\mathbf{C}' = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{A}'\mathbf{C}' = \begin{bmatrix} -1 \\ -3 \\ -1 \end{bmatrix}, \quad (\mathbf{A}')^2\mathbf{C}' = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix}$$

note que

$$\begin{vmatrix} 1 & -1 & 0 \\ 1 & -3 & 5 \\ 0 & -1 & 0 \end{vmatrix} = 5$$

Portanto a característica da matriz $[\mathbf{C}'|\mathbf{A}'\mathbf{C}'|(\mathbf{A}')^2\mathbf{C}']$ é três, e o sistema é completamente observável.

Problema A.16.2 Mostre que o sistema de tempo contínuo descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [u] \tag{16.51}$$

$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{16.52}$$

é de estado completamente controlável e completamente observável.

O sistema de tempo discreto, que é equivalente ao sistema de tempo contínuo das Eqs. (16.51) e (16.52), é dado por

$$\begin{bmatrix} x_1((k+1)T) \\ x_2((k+1)T) \end{bmatrix} = \begin{bmatrix} \cos T & \operatorname{sen} T \\ -\operatorname{sen} T & \cos T \end{bmatrix} \begin{bmatrix} x_1(kT) \\ x_2(kT) \end{bmatrix} + \begin{bmatrix} 1 - \cos T \\ \operatorname{sen} T \end{bmatrix} [u(kT)]$$

$$y(kT) = [1 \quad 0] \begin{bmatrix} x_1(kT) \\ x_2(kT) \end{bmatrix}$$

onde T é o período de amostragem. Mostre que o sistema de tempo discreto é de estado completamente controlável e completamente observável se e apenas se $T \neq n\pi$ $(n = 0, 1, 2, \ldots)$.

Solução. Para o sistema de tempo contínuo descrito pelas Eqs. (16.51) e (16.52),

$$[\mathbf{B} \mid \mathbf{AB}] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

A característica de $[\mathbf{B}|\mathbf{AB}]$ é 2 e o sistema é de estado completamente controlável. Agora,

$$[\mathbf{C}' \mid \mathbf{A}'\mathbf{C}'] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

A característica de $[\mathbf{C}'|\mathbf{A}'\mathbf{C}']$ também é 2 e o sistema é completamente observável.

Para o sistema de tempo discreto, temos

$$[\mathbf{H} \mid \mathbf{GH}] = \begin{bmatrix} 1 - \cos T & 1 + \cos T - 2\cos^2 T \\ \operatorname{sen} T & -\operatorname{sen} T + 2\cos T \operatorname{sen} T \end{bmatrix}$$

A característica da matriz $[\mathbf{H}|\mathbf{GH}]$ é 2 se e apenas se $T \neq n\pi$ $(n = 0, 1, 2, \ldots)$. A característica de

$$[\mathbf{C}' \mid \mathbf{G}'\mathbf{C}'] = \begin{bmatrix} 1 & \cos T \\ 0 & \operatorname{sen} T \end{bmatrix}$$

é 2 se e apenas se $T \neq n\pi (n = 0,1,2,\ldots)$. Portanto, o sistema de tempo discreto é de estado completamente controlável e completamente observável se e apenas se $T \neq n\pi$ $(n = 0, 1, 2, \ldots)$.

É importante ressaltar que o processo de amostragem piora a controlabilidade e a observabilidade. (Fisicamente, no caso presente o sinal de saída contínuo é um sinal senoidal, mas o sinal de saída amostrado é zero se $T = n\pi$.) Note que podemos evitar sempre a perda da controlabilidade e observabilidade pela escolha de períodos de amostragem suficientemente pequenos quando comparados com a menor constante de tempo do sistema.

Problema A.16.3 Considere o sistema de tempo discreto discutido no Exemplo 16.9. Vimos que qualquer estado inicial pode ser conduzido à origem no máximo em dois períodos de amostragem se a amplitude de $u(kT)$ não for limitada. Entretanto, se a amplitude de $u(kT)$ é limitada, então alguns estados iniciais não podem ser transferidos para a origem em dois períodos de amostragem. Eles podem requerer três, quatro ou mais períodos de amostragem.

Suponha que a amplitude de $u(kT)$ é limitada, ou

$$|u(kT)| \leq 1$$

Determine as regiões de estados iniciais no plano x_1 e x_2 que podem ser transferidas para a origem em um período de amostragem e dois períodos de amostragem, respectivamente. Suponha que $T = 1$ s.

Solução. Se impusermos $\mathbf{x}(T) = \mathbf{0}$, então, da Eq. (16.31), temos

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 0{,}72 \\ -1{,}72 \end{bmatrix} [u(0)]$$

Como $|u(0)| \leq 1$, obtemos

$$|x_1(0)| \leq 0{,}72, \qquad |x_2(0)| \leq 1{,}72$$

Portanto, se o estado inicial estiver sobre o segmento de reta

$$1{,}72x_1(0) + 0{,}72x_2(0) = 0, \qquad -0{,}72 \leq x_1(0) \leq 0{,}72$$

ele pode ser levado para a origem em um período de amostragem, ou 1 s. Este segmento de reta é visto na Fig. 16.15.

Se impusermos $\mathbf{x}(2) = \mathbf{0}$, então, da Eq. (16.30), obtemos

$$u(0) = -1{,}58x_1(0) - 1{,}24x_2(0)$$
$$u(1) = 0{,}58x_1(0) + 0{,}24x_2(0)$$

Como $|u(0)| \leq 1$ e $|u(1)| \leq 1$, obtemos as seguintes quatro relações:

$$1{,}58x_1(0) + 1{,}24x_2(0) \leq 1 \tag{16.53}$$

$$1{,}58x_1(0) + 1{,}24x_2(0) \geq -1 \tag{16.54}$$

$$0{,}58x_1(0) + 0{,}24x_2(0) \leq 1 \tag{16.55}$$

$$0{,}58x_1(0) + 0{,}24x_2(0) \geq -1 \tag{16.56}$$

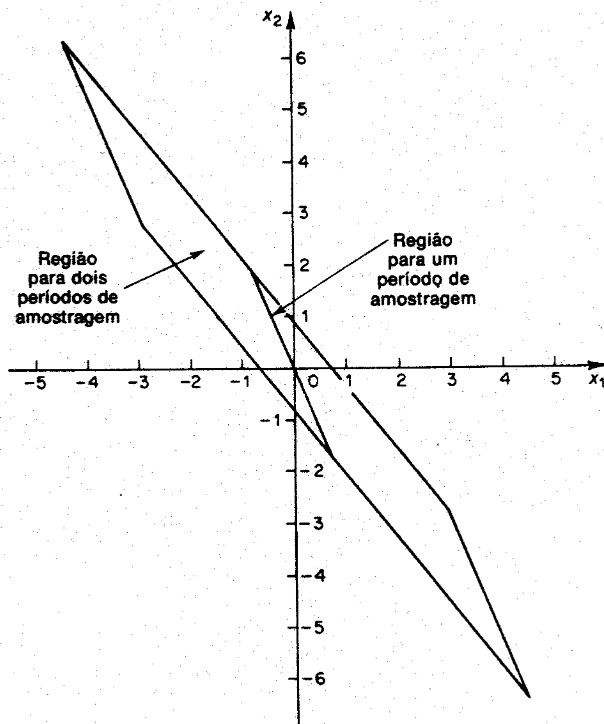


Fig. 16.15 Regiões a partir das quais estados iniciais podem ser levados à origem em um ou dois períodos de amostragem.

A região limitada pelas Eqs. (16.53) a (16.56) é vista na Fig. 16.15. Se o estado inicial está dentro desta região, então ele pode ser transferido para a origem em dois períodos de amostragem, ou 2 s.

Problema A.16.4 Considere o sistema de controle de tempo ótimo visto na Fig. 16.2. Referindo-se à Fig. 16.4, obtenha a equação para a curva de chaveamento AOB. Se o torque máximo é aumentado, como se altera a forma da curva de chaveamento?

Supondo que $T/J = 2$ e que a condição inicial está no primeiro quadrante do plano e—$\dot{e}$, faça um gráfico de curvas típicas de e *versus* t, e u *versus* t.

Solução. A equação da curva AO na Fig. 16.4 é

$$\dot{e}^2 + 2\frac{T}{J}e = 0$$

A equação da curva BO é

$$\dot{e}^2 - 2\frac{T}{J}e = 0$$

A equação da curva de chaveamento pode ser escrita combinando-se as duas equações anteriores, como segue:

$$|\dot{e}|\,\dot{e} + 2\frac{T}{J}e = 0$$

A mudança na forma da curva de chaveamento quando o torque máximo é aumentado pode ser descrita como segue: Quanto maior o torque máximo, mais perto está a curva de chaveamento do eixo de ordenada no plano de fase. A Fig. 16.16 mostra duas curvas de chaveamento para dois valores de T.

Finalmente, a Fig. 16.17 mostra uma trajetória típica de tempo ótimo no plano e—$\dot{e}$ e as correspondentes curvas de e *versus* t e u *versus* t.

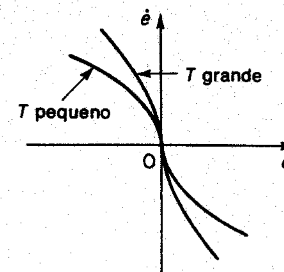


Fig. 16.16 Curvas de chaveamento para dois valores do torque máximo T.

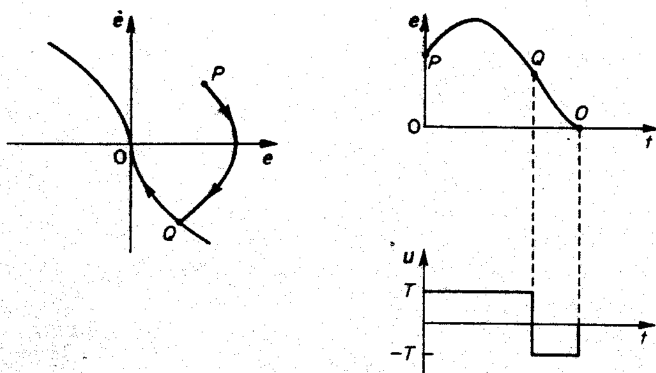


Fig. 16.17 Trajetória típica de tempo ótimo no plano *e—ė* e as correspondentes curvas de *e versus t* e *u versus t*.

Problema A.16.5 A curva de chaveamento para o sistema de controle ótimo visto na Fig. 16.2 é a parábola *AOB*, como visto na Fig. 16.4. O chaveamento de torque positivo para negativo pode ser conseguido através de um computador especializado. Entretanto, uma aproximação linear para a curva de chaveamento dá, em muitos casos, um resultado tão bom que não há muita justificativa para complicar o controlador introduzindo dispositivos de computação a fim de determinar o chaveamento ótimo. Portanto, em muitos casos práticos, pode ser suficiente se ter chaveamentos ocorrendo ao longo de uma linha reta ao invés de uma parábola.

Considere o sistema visto na Fig. 16.18. Faça o gráfico da curva de chaveamento e de uma trajetória típica.

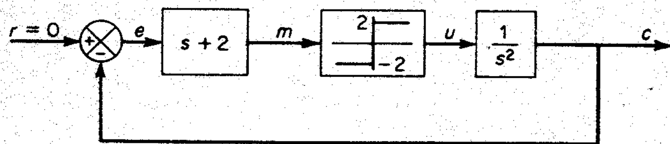


Fig. 16.18 Sistema de controle.

Solução. Do diagrama de bloco, a equação da linha de chaveamento *AOB* pode ser escrita como

$$\dot{e} + 2e = 0$$

A equação que descreve o processo é

$$\ddot{c} = u = 2 \qquad \text{para } \dot{e} + 2e > 0$$
$$= -2 \qquad \text{para } \dot{e} + 2e < 0$$

Notando que $r = 0$, temos $c = -e$. Portanto,

$$\ddot{e} = -2 \qquad \text{para } \dot{e} + 2e > 0$$
$$= 2 \qquad \text{para } \dot{e} + 2e < 0$$

A equação para as trajetórias é

$$\dot{e}^2 = \pm 4e + K$$

onde K é uma constante. Esta equação descreve duas famílias de parábolas no plano $e - \dot{e}$. A Fig. 16.19 mostra a linha de chaveamento e uma trajetória típica. Note que alguém poderia pensar que a trajetória pára no ponto P. Na prática, entretanto, devido à existência de pequenos atrasos de tempo no chaveamento, ocorrem desvios na linha de chaveamento, e a trajetória converge para a origem do plano de fase como visto na Fig. 16.19. (Um sistema com tal linha reta de chaveamento não é de tempo ótimo. Mas o tempo de resposta é razoavelmente curto e adequado para as finalidades práticas.)

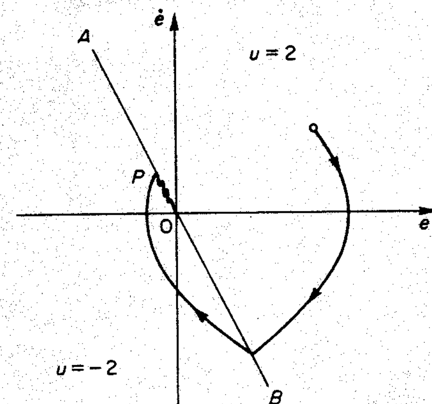


Fig. 16.19 Linha de chaveamento e uma trajetória típica.

Problema A.16.16 Considere o sistema

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

onde **x** e **u** são vetores reais e **A** e **B** matrizes reais. Mostre que se o vetor de controle **u** é dado por

$$\mathbf{u} = -\mathbf{B}'\mathbf{P}\mathbf{x} \qquad (16.57)$$

onde **P** é uma matriz simétrica real positiva definida satisfazendo a condição de que

$$\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{I}$$

então a origem do sistema é assintoticamente estável globalmente.

Solução. Vamos escolher o seguinte V como uma possível função de Liapunov:

$$V = \mathbf{x}'\mathbf{P}\mathbf{x}$$

Então

$$\begin{aligned}\dot{V} &= \dot{\mathbf{x}}'\mathbf{P}\mathbf{x} + \mathbf{x}'\mathbf{P}\dot{\mathbf{x}} \\ &= \mathbf{x}'(\mathbf{A}'\mathbf{P} + \mathbf{P}\mathbf{A})\mathbf{x} + \mathbf{u}'\mathbf{B}'\mathbf{P}\mathbf{x} + \mathbf{x}'\mathbf{P}\mathbf{B}\mathbf{u} \qquad (16.58) \\ &= -\mathbf{x}'\mathbf{I}\mathbf{x} + 2\mathbf{x}'\mathbf{P}\mathbf{B}\mathbf{u}\end{aligned}$$

Se a Eq. (16.57) é substituída na Eq. (16.58), vemos que

$$\dot{V} = -\mathbf{x}'\mathbf{I}\mathbf{x} - 2\mathbf{x}'\mathbf{P}\mathbf{B}\mathbf{B}'\mathbf{P}\mathbf{x}$$

Como **PBB'P** sempre é positiva definida ou positiva semidefinida, $\dot{V}$ sempre é negativa definida. Como V é positiva definida, $\dot{V}$ é negativa definida, e $\dot{V} \to \infty$ para $\|\mathbf{x}\| \to \infty$, a origem do sistema é assintoticamente estável globalmente.

Problema A.16.7 Considere o sistema de controle descrito por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \qquad (15.59)$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Supondo uma lei de controle linear

$$u = -\mathbf{K}\mathbf{x} = -k_1x_1 - k_2x_2 \qquad (16.60)$$

determine as constantes k_1 e k_2 de tal forma que o seguinte índice de desempenho é minimizado:

$$J = \int_0^\infty \mathbf{x}'\mathbf{x}\, dt$$

Considere apenas o caso onde a condição inicial é

$$\mathbf{x}(0) = \begin{bmatrix} c \\ 0 \end{bmatrix}$$

Escolha a freqüência natural não amortecida como 2 rad/s.

Solução. Substituindo a Eq. (16.60) na Eq. (16.59), obtemos

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} - \mathbf{B}\mathbf{K}\mathbf{x}$$

ou

$$\begin{aligned}\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}[-k_1x_1 - k_2x_2] \\ &= \begin{bmatrix} 0 & 1 \\ -k_1 & -k_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad (16.61)\end{aligned}$$

Portanto

$$\mathbf{A} - \mathbf{B}\mathbf{K} = \begin{bmatrix} 0 & 1 \\ -k_1 & -k_2 \end{bmatrix}$$

A eliminação de x_2 da Eq. (16.61) nos dá

$$\ddot{x}_1 + k_2\dot{x}_1 + k_1x_1 = 0$$

Como a freqüência natural não amortecida é especificada como 2 rad/s, obtemos

$$k_1 = 4$$

Portanto

$$\mathbf{A} - \mathbf{B}\mathbf{K} = \begin{bmatrix} 0 & 1 \\ -4 & -k_2 \end{bmatrix}$$

A − BK é uma matriz estável se $k_2 > 0$. Nosso problema agora é o de determinar o valor de k_2 de tal forma que o índice de desempenho

$$J = \int_0^\infty \mathbf{x}'\mathbf{x}\, dt = \mathbf{x}'(0)\mathbf{P}\mathbf{x}(0)$$

é minimizado, onde a matriz **P** é determinada de

$$(\mathbf{A} - \mathbf{B}\mathbf{K})'\mathbf{P} + \mathbf{P}(\mathbf{A} - \mathbf{B}\mathbf{K}) = -\mathbf{I}$$

ou

$$\begin{bmatrix} 0 & -4 \\ 1 & -k_2 \end{bmatrix}\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -4 & -k_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Resolvendo para a matriz **P**, obtemos

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} \frac{5}{2k_2} + \frac{k_2}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{5}{8k_2} \end{bmatrix}$$

O índice de desempenho é então

$$\begin{aligned}J &= \mathbf{x}'(0)\mathbf{P}\mathbf{x}(0) \\ &= [c \quad 0]\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}\begin{bmatrix} c \\ 0 \end{bmatrix} = p_{11}c^2 \\ &= \left(\frac{5}{2k_2} + \frac{k_2}{8}\right)c^2 \qquad (16.62)\end{aligned}$$

Para minimizar J, diferenciamos J com respeito a k_2 e o igualamos a zero como segue:

$$\frac{\partial J}{\partial k_2} = \left(\frac{-5}{2k_2^2} + \frac{1}{8}\right)c^2 = 0$$

Portanto

$$k_2 = \sqrt{20}$$

Com este valor de k_2 temos $\partial^2 J/\partial k_2^2 > 0$. Portanto o valor mínimo de J é obtido substituindo-se $k_2 = \sqrt{20}$ na Eq. (16.62), ou

$$J_{\text{mín.}} = \frac{\sqrt{5}}{2}c^2$$

O sistema projetado tem a lei de controle

$$u = -4x_1 - \sqrt{20}x_2$$

O sistema projetado é ótimo no sentido de que ele resulta em um valor mínimo para o índice de desempenho J sob a condição inicial suposta.

Problema A.16.8 Considere o sistema de controle visto na Fig. 16.9. A equação para o processo é

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \tag{16.63}$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Supondo a lei de controle linear

$$u = -\mathbf{K}\mathbf{x} = -k_1x_1 - k_2x_2 \tag{16.64}$$

determine as constantes k_1 e k_2 de tal forma que o seguinte índice de desempenho é minimizado:

$$J = \int_0^\infty (\mathbf{x}'\mathbf{x} + u'u)\, dt$$

Solução. Substituindo a Eq. (16.64) na Eq. (16.63), obtemos

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} - \mathbf{B}\mathbf{K}\mathbf{x} = (\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{x} = \begin{bmatrix} 0 & 1 \\ -k_1 & -k_2 \end{bmatrix}\mathbf{x}$$

Se supusermos k_1 e k_2 constantes positivas, então $\mathbf{A} - \mathbf{B}\mathbf{K}$ se torna uma matriz estável e $\mathbf{x}(\infty) = \mathbf{0}$. Portanto o índice de desempenho pode ser escrito como

$$\begin{aligned} J &= \int_0^\infty (\mathbf{x}'\mathbf{x} + \mathbf{x}'\mathbf{K}'\mathbf{K}\mathbf{x})\, dt \\ &= \int_0^\infty \mathbf{x}'(\mathbf{I} + \mathbf{K}'\mathbf{K})\mathbf{x}\, dt \\ &= \mathbf{x}'(0)\mathbf{P}\mathbf{x}(0) \end{aligned}$$

onde $\mathbf{P}$ é determinada de

$$(\mathbf{A} - \mathbf{B}\mathbf{K})'\mathbf{P} + \mathbf{P}(\mathbf{A} - \mathbf{B}\mathbf{K}) = -(\mathbf{I} + \mathbf{K}'\mathbf{K})$$

ou

$$\begin{bmatrix} 0 & -k_1 \\ 1 & -k_2 \end{bmatrix}\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -k_1 & -k_2 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} k_1^2 & k_1k_2 \\ k_1k_2 & k_2^2 \end{bmatrix}$$

Esta equação matricial resulta nas seguintes três equações em p_{ij}:

$$\begin{aligned} -2k_1p_{12} &= -1 - k_1^2 \\ p_{11} - k_2p_{12} - k_1p_{22} &= -k_1k_2 \\ 2p_{12} - 2k_2p_{22} &= -1 - k_2^2 \end{aligned}$$

Resolvendo estas três equações para obter os p_{ij}, temos

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\left(\frac{k_2}{k_1} + \frac{k_1}{k_2}\right) + \frac{k_1}{2k_2}\left(\frac{1}{k_1} + k_1\right) & \frac{1}{2}\left(\frac{1}{k_1} + k_1\right) \\ \frac{1}{2}\left(\frac{1}{k_1} + k_1\right) & \frac{1}{2}\left(\frac{1}{k_2} + k_2\right) + \frac{1}{2k_2}\left(\frac{1}{k_1} + k_1\right) \end{bmatrix}$$

Agora

$$\begin{aligned} J &= \mathbf{x}'(0)\mathbf{P}\mathbf{x}(0) \\ &= \left[\frac{1}{2}\left(\frac{k_2}{k_1} + \frac{k_1}{k_2}\right) + \frac{k_1}{2k_2}\left(\frac{1}{k_1} + k_1\right)\right]x_1^2(0) + \left(\frac{1}{k_1} + k_1\right)x_1(0)x_2(0) \\ &\quad + \left[\frac{1}{2}\left(\frac{1}{k_2} + k_2\right) + \frac{1}{2k_2}\left(\frac{1}{k_1} + k_1\right)\right]x_2^2(0) \end{aligned}$$

Para minimizar J, fazemos $\partial J/\partial k_1 = 0$ e $\partial J/\partial k_2 = 0$, ou

$$\frac{\partial J}{\partial k_1} = \left[\frac{1}{2}\left(\frac{-k_2}{k_1^2} + \frac{1}{k_2}\right) + \frac{k_1}{k_2}\right]x_1^2(0) + \left(\frac{-1}{k_1^2} + 1\right)x_1(0)x_2(0) + \left[\frac{1}{2k_2}\left(\frac{-1}{k_1^2} + 1\right)\right]x_2^2(0) = 0$$

$$\frac{\partial J}{\partial k_2} = \left[\frac{1}{2}\left(\frac{1}{k_1} - \frac{k_1}{k_2^2}\right) + \frac{-k_1}{2k_2^2}\left(\frac{1}{k_1} + k_1\right)\right]x_1^2(0) + \left[\frac{1}{2}\left(\frac{-1}{k_2^2} + 1\right) - \frac{1}{2k_2^2}\left(\frac{1}{k_1} + k_1\right)\right]x_2^2(0) = 0$$

Para quaisquer condições iniciais $x_1(0)$ e $x_2(0)$, o valor de J se torna mínimo quando

$$k_1 = 1, \qquad k_2 = \sqrt{3}$$

Note que k_1 e k_2 são constantes positivas como supusemos na solução. Portanto, para a lei de controle ótimo,

$$\mathbf{K} = [k_1 \quad k_2] = [1 \quad \sqrt{3}]$$

O diagrama de bloco deste sistema de controle ótimo é visto na Fig. 16.10 com $\mu = 1$.

Problema A.16.9 Considere o sistema

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$$

que pode ser linear ou não-linear. Deseja-se determinar a lei de controle ótimo $\mathbf{u} = \mathbf{g}(\mathbf{x})$ tal que o seguinte índice de desempenho,

$$J = \int_0^\infty L(\mathbf{x}, \mathbf{u})\, dt$$

é minimizado, onde $\mathbf{u}$ não é limitado.

Se a origem do sistema descrito por

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{g}(\mathbf{x}))$$

é assintoticamente estável, e portanto existe uma função de Liapunov $V(\mathbf{x})$ tal que $\dot{V}(\mathbf{x})$ é negativa definida, então mostre que uma condição suficiente para um vetor de controle $\mathbf{u}_1$ ser ótimo é que $H(\mathbf{x}, \mathbf{u})$, onde

$$H(\mathbf{x}, \mathbf{u}) = \frac{dV}{dt} + L(\mathbf{x}, \mathbf{u}) \tag{16.65}$$

é mínimo com $\mathbf{u} = \mathbf{u}_1$, ou

$$\begin{aligned} \min_{\mathbf{u}} H(\mathbf{x}, \mathbf{u}) &= \min_{\mathbf{u}} \left[\frac{dV}{dt} + L(\mathbf{x}, \mathbf{u})\right] \\ &= \left.\frac{dV}{dt}\right|_{\mathbf{u}=\mathbf{u}_1} + L(\mathbf{x}, \mathbf{u}_1) \end{aligned} \tag{16.66}$$

e

$$\left.\frac{dV}{dt}\right|_{\mathbf{u}=\mathbf{u}_1} = -L(\mathbf{x}, \mathbf{u}_1) \tag{16.67}$$

Solução. Vamos integrar ambos os lados da Eq. (16.67). Então,

$$V(\mathbf{x}(\infty)) - V(\mathbf{x}(0)) = -\int_0^\infty L(\mathbf{x}(t), \mathbf{u}_1(t))\, dt \tag{16.68}$$

Como a origem do sistema é assintoticamente estável, $\mathbf{x}(\infty) = \mathbf{0}$ e $V(\mathbf{x}(\infty)) = 0$. Então a Eq. (16.68) se torna

$$V(\mathbf{x}(0)) = \int_0^\infty L(\mathbf{x}(t), \mathbf{u}_1(t))\, dt \tag{16.69}$$

Para provar que $\mathbf{u}_1(t)$ é ótimo, suponha que $\mathbf{u}_1(t)$ não é ótimo e que o vetor de controle $\mathbf{u}_2(t)$ nos dará um valor pequeno de J. Então,

$$\int_0^\infty L(\mathbf{x}(t), \mathbf{u}_2(t))\, dt < \int_0^\infty L(\mathbf{x}(t), \mathbf{u}_1(t))\, dt$$

Note que da Eq. (16.66), o valor mínimo de $H(\mathbf{x}, \mathbf{u})$ ocorre em $\mathbf{u} = \mathbf{u}_1$. Note também que, da Eq. (16.67), este valor mínimo é igual a zero. Portanto,

$$H(\mathbf{x}, \mathbf{u}) \geq 0$$

para todo $\mathbf{u}$. Assim,

$$H(\mathbf{x}, \mathbf{u}_2) = \left.\frac{dV}{dt}\right|_{\mathbf{u}=\mathbf{u}_2} + L(\mathbf{x}, \mathbf{u}_2) \geq 0$$

Integrando ambos os lados desta desigualdade de 0 a ∞, obtemos

$$V(\mathbf{x}(\infty)) - V(\mathbf{x}(0)) \geq -\int_0^\infty L(\mathbf{x}(t), \mathbf{u}_2(t))\, dt$$

Como $V(\mathbf{x}(\infty)) = 0$, temos

$$V(\mathbf{x}(0)) \leq \int_0^\infty L(\mathbf{x}(t), \mathbf{u}_2(t))\, dt \tag{16.70}$$

Então, das Eqs. (16.69) e (16.70),

$$\int_0^\infty L(\mathbf{x}(t), \mathbf{u}_1(t))\, dt \leq \int_0^\infty L(\mathbf{x}(t), \mathbf{u}_2(t))\, dt$$

Isto é uma contradição. Portanto $\mathbf{u}_1(t)$ é o vetor de controle ótimo.

Problema A.16.10 Considere o sistema

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [u]$$

Deseja-se achar a função de controle ótimo u tal que o seguinte índice de desempenho,

$$J = \int_0^\infty (\mathbf{x}'\mathbf{Q}\mathbf{x} + u'u)\, dt, \qquad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix}$$

seja minimizado.

Referindo-se à condição suficiente para o vetor de controle ótimo apresentada no Problema A.16.9, determine o sinal de controle ótimo $u(t)$.

Solução. Vamos definir

$$V = \mathbf{x}'\mathbf{P}\mathbf{x}$$

onde

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}$$

Então

$$V = p_{11}x_1^2 + 2p_{12}x_1x_2 + p_{22}x_2^2$$

e

$$\dot{V} = 2p_{11}x_1x_2 + 2p_{12}x_2^2 + (2p_{12}x_1 + 2p_{22}x_2)u$$

A Eq. (16.65) se torna

$$H(\mathbf{x}, u) = \frac{dV}{dt} + L(\mathbf{x}, u) \tag{16.71}$$

$$= 2p_{11}x_1x_2 + 2p_{12}x_2^2 + (2p_{12}x_1 + 2p_{22}x_2)u + x_1^2 + \mu x_2^2 + u^2$$

O sinal de controle ótimo $u(t)$ é aquele que minimiza $H(\mathbf{x}, u)$. Portanto, diferenciamos $H(\mathbf{x}, u)$ com relação a u e igualamos a zero, ou

$$\frac{\partial H}{\partial u} = 2p_{12}x_1 + 2p_{22}x_2 + 2u = 0$$

de que obtemos

$$u = -p_{12}x_1 - p_{22}x_2 \tag{16.72}$$

Substituindo a Eq. (16.72) na Eq. (16.71) e igualando o resultado a zero, obtemos

$$x_1^2(1 - p_{12}^2) + x_1x_2(2p_{11} - 2p_{12}p_{22}) + x_2^2(\mu + 2p_{12} - p_{22}^2) = 0$$

Esta última equação deve valer para qualquer x_1 e x_2. Portanto, impomos

$$1 - p_{12}^2 = 0$$
$$p_{11} - p_{12}p_{22} = 0$$
$$\mu + 2p_{12} - p_{22}^2 = 0$$

(Note que estas três equações em p_{ij} são idênticas àquelas obtidas no Exemplo 16.12.) Resolvendo estas três equações simultâneas para p_{11}, p_{12} e p_{22}, impondo que **P** seja positiva definida, obtemos

$$p_{11} = \sqrt{\mu + 2}, \qquad p_{12} = 1, \qquad p_{22} = \sqrt{\mu + 2}$$

A lei de controle ótimo é então dada por

$$u = -x_1 - \sqrt{\mu + 2}x_2$$

Com esta lei, a equação de estado do sistema se torna

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [-x_1 - \sqrt{\mu + 2}x_2]$$

ou

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -\sqrt{\mu + 2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

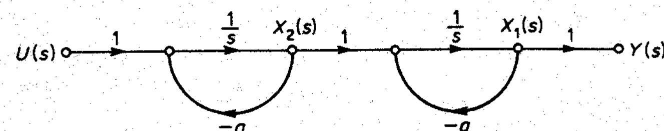
Note que a matriz de coeficientes nesta última equação é estável. Portanto a origem do sistema é assintoticamente estável e a presente abordagem para a solução é válida.

PROBLEMAS

Problema B.16.1 Determine a controlabilidade de estado e a observabilidade do sistema descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} [u]$$

$$y = [4 \quad 5 \quad 1] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$



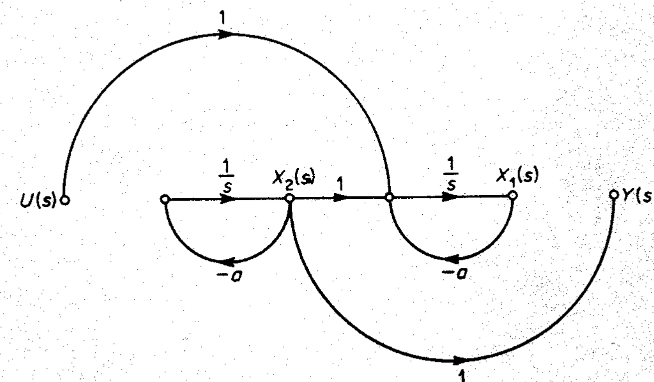


Fig. 16.20 Diagramas de fluxo de sinais de sistemas.

Problema B.16.2 São os sistemas vistos nas Figs. 16.20(a) e (b) controláveis e observáveis?

Problema B.16.3 Considere o sistema descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} a \\ b \\ c \end{bmatrix} [u]$$

Determine as condições em a, b, e c para controlabilidade completa de estado.

Problema B.16.4 Considere o sistema descrito por

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} [u]$$

$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Determine as condições em a, b, c e d para controlabilidade completa de estado e observabilidade completa.

Problema B.16.5 Considere o sistema

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} [u]$$

$$y = [1 \quad 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Mostre que a saída y é independente da função de controle u.

Problema B.16.6 Suponha que um sistema mecânico vibratório é descrito por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

onde

$\mathbf{x}$ = vetor de estado (vetor n-dimensional)
$\mathbf{u}$ = vetor de entrada (vetor r-dimensional)
$\mathbf{A}$ = matriz constante $n \times n$
$\mathbf{B}$ = matriz constante $n \times r$

Determine a condição em **A** e **B** de tal forma que todos os modos de oscilação podem ser excitados.

Problema B.16.7 Considere o sistema de controle de tempo ótimo visto na Fig. 16.21. Obtenha a equação da curva de chaveamento. Faça um gráfico da curva de chaveamento no plano e—$\dot{e}$ e mostre algumas trajetórias de tempo ótimo típicas.

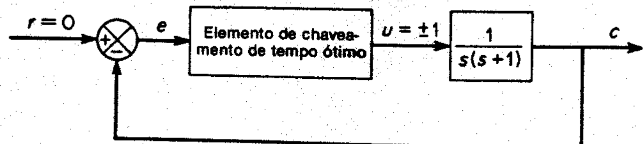


Fig. 16.21 Sistema de controle de tempo ótimo.

Problema B.16.8 Faça um gráfico de plano de fase do sistema visto na Fig. 16.22.

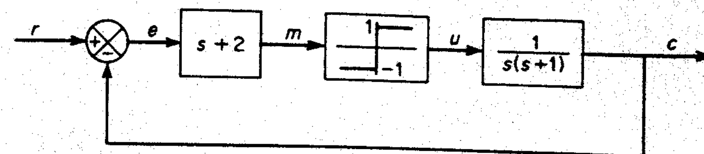


Fig. 16.22 Sistema de controle.

Problema B.16.9 Considere o sistema descrito por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \qquad \mathbf{x}(0) = \mathbf{c}$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -2 & -a \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c_1 \\ 0 \\ 0 \end{bmatrix}$$

$$a = \text{parâmetro ajustável} > 0$$

Determine o valor do parâmetro a de tal forma que minimize o seguinte índice de desempenho:

$$J = \int_0^\infty \mathbf{x}'\mathbf{x}\, dt$$

Problema B.16.10 Considere o sistema visto na Fig. 16.23. Determine o valor do ganho K de tal forma que o coeficiente de amortecimento ζ do sistema a malha-fechada seja igual a 0,5. Em seguida calcule a freqüência natural não-amortecida ω_n do sistema a malha-fechada. Supondo que $e(0) = 1$ e que $\dot{e}(0) = 0$, calcule

$$\int_0^\infty e^2(t)\, dt$$

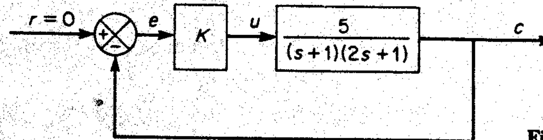


Fig. 16.23 Sistema de controle.

Problema B.16.11 Determine a função de controle ótimo u para o sistema descrito por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u$$

onde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

tal que o seguinte índice de desempenho seja minimizado:

$$J = \int_0^\infty (\mathbf{x}'\mathbf{x} + u'u)\, dt$$

Referências

A-1 Andronow, A. A., and C. E. Chaikin, *Theory of Oscillations*, Princeton, N.J.: Princeton University Press, 1949.

A-2 Aseltine, J. A., *Transform Methods in Linear System Analysis*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1958.

A-3 Athanassiades, M., and P. L. Falb, "Time Optimal Control for Plants with Numerator Dynamics," *IRE Trans. Automatic Control* (1962), pp. 46–7.

A-4 Athans, M., and P. L. Falb, *Optimal Control: An Introduction to the Theory and Its Applications*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1965.

B-1 Bayliss, L. E., *Living Control Systems*, London, England: English Universities Press Limited, 1966.

B-2 Bellman, R., I. Glicksberg, and O. Gross, "On the Bang-Bang Control Problem," *Quart. Appl. Math.*, **14** (1956), pp. 11–8.

B-3 Brown, G. S., and D. P. Campbell, *Principles of Servomechanisms*, New York, N.Y.: John Wiley & Sons, Inc., 1955.

B-4 Buland, R. N., "Analysis of Nonlinear Servos by Phase Plane-Delta Method," *J. Franklin Inst.*, 257 (1954), pp. 37–48.

B-5 Bushaw, D. W., "Differential Equations with a Discontinuous Forcing Term," *Stevens Inst. Tech. Experimental Towing Tank Report* **469**, Hoboken, N.J., Jan., 1953.

B-6 Bushaw, D. W., "Optimal Discontinuous Forcing Terms," in S. Lefschetz (ed.), *Contributions to the Theory of Nonlinear Oscillations*, 4, Princeton, N.J.: Princeton University Press, 1958, pp. 29–52.

B-7 Butman, S., and R. Sivan (Sussman), "On Cancellations, Controllability and Observability," *IEEE Trans. Automatic Control* (1964), pp. 317–8.

C-1 Campbell, D. P., *Process Dynamics*, New York, N.Y.: John Wiley & Sons, Inc., 1958.

C-2 Chandaket, P., C. T. Leondes, and E. C. Deland, "Optimum Non-linear Bang-Bang Control Systems with Complex Roots," *AIEE Trans. Part II*, **80** (1961), pp. 82–102.

C-3 Chang, S. S. L., *Synthesis of Optimum Control Systems*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1961.

C-4 Cheng, D. K., *Analysis of Linear Systems*, Reading, Mass.: Addison-Wesley Publishing Company, Inc., 1959.

C-5 Chestnut, H., and R. W. Mayer, *Servomechanisms and Regulating System Design*, vol. 1, New York, N.Y.: John Wiley & Sons, Inc., 1959.

C-6 Coddington, E. A., and N. Levinson, *Theory of Ordinary Differential Equations*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1955.

C-7 Cosgriff, R. L., *Nonlinear Control Systems*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1958.

C-8 Cunningham, W. J., *Introduction to Nonlinear Analysis*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1958.

D-1 D'Azzo, J. J., and C. H. Houpis, *Feedback Control System Analysis and Synthesis*, 2nd ed., New York, N.Y.: McGraw-Hill Book Company, Inc., 1966.

D-2 Del Toro, V., and S. R. Parker, *Principles of Control System Engineering*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1960.

D-3 DeRusso, P. M., R. J. Roy, and C. M. Close, *State Variables for Engineers*, New Yrok, N.Y.: John Wiley & Sons, Inc., 1965.

D-4 Dorf, R,, *Time Domain Analysis and Design of Control Systems*, Reading, Mass.: Addison-Wesley Publishing Company, Inc., 1965.

E-1 Eckman, D. P., *Automatic Process Control*, New York, N.Y.: John Wiley & Sons, Inc., 1958.

E-2 Enns, M., J. R. Greenwood, III, J. E. Matheson, and F. T. Thompson, "Practical Aspects of State-Space Methods Part I: System Formulation and Reduction," *IEEE Trans. Military Electronics* (1964), pp. 81–93.

E-3 Evans, W. R., "Graphical Analysis of Control Systems," *AIEE Trans. Part II*, **67** (1948), pp. 547–51.

E-4 Evans, W. R., "Control System Synthesis by Root Locus Method," *AIEE Trans. Part II*, **69** (1950), pp. 66–9.

E-5 Evans, W. R., "The Use of Zeros and Poles for Frequency Response or Transient Response," *ASME Trans.*, **76** (1954), pp. 1335–44.

F-1 Flügge-Lotz, I., "Synthesis of Third-order Contactor Control Systems," *Proc. First Intern, Cong. IFAC*, Moscow, 1960; *Automatic and Remote*

Control, London, England: Butterworths & Company, Ltd., 1961, pp. 390–7.

F-2 Flügge-Lotz, I., and T. Ishikawa, "Investigation of Third-order Contactor Control Systems with Two Complex Poles without Zeros," *NASA Tech. Note D248*, 1960.

F-3 Freeman, H., *Discrete-Time Systems*, New York, N.Y.: John Wiley & Sons, Inc., 1965.

F-4 Friedland, B., "The Structure of Optimum Control Systems," *ASME J. Basic Engineering*, ser. D, **84** (1962), pp. 1–12.

F-5 Fuller, A. T., "Phase Space in the Theory of Optimum Control," *J. Elec. Control*, ser. 1 VIII (1960), pp. 381–400.

G-1 Gardener, M. F., and J. L. Barnes, *Transients in Linear Systems*, vol 1, New York, N.Y.: John Wiley & Sons, Inc., 1942.

G-2 Gibson, J. E., *Nonlinear Automatic Control*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1963.

G-3 Gilbert, E. G., "Controllability and Observability in Multivariable Control Systems," *J. SIAM Control*, ser. A, **1** (1963), pp. 128–51.

G-4 Gille, J. C., M. J. Pelegrin, and P. Decaulne, *Feedback Control Systems*, New York, N.Y.: McGraw-Hill Book Company, 1959.

G-5 Graham, D., and R. C. Lathrop, "The Synthesis of Optimum Response: Criteria and Standard Forms," *AIEE Trans. Part II*, **72** (1953), pp. 273–88.

G-6 Graham, D., and D. McRuer, *Analysis of Nonlinear Control Systems*, New York, N.Y.: John Wiley & Sons, Inc., 1961.

G-7 Gupta, S. C., *Transform and State Variable Methods in Linear Systems*, New York, N.Y.: John Wiley & Sons, Inc., 1966.

H-1 Hahn, W., *Theory and Application of Liapunov's Direct Method*, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963.

H-2 Harrison, H. L., and J. G. Bollinger, *Introduction to Automatic Controls*, Scranton, Pa.: International Textbook Company, Inc., 1963.

H-3 Harvey, C. A., "Determining the Switching Criterion for Time Optimal Control," *J. Math. Anal. Appl.*, **5** (1962), pp. 245–57.

H-4 Harvey, C. A., and E. B. Lee, "On the Uniqueness of Time-optimal Control for Linear Processes," *J. Math . Anal. Appl.*, **5** (1962), pp. 258–68.

H-5 Henke, R. W., "Digital Fluidics Works Now," Control Engineering, **14**, no. 1, Jan., 1967, pp. 100–4.

H-6 Higdon, D. T., and R. H. Cannon, Jr., "On the Control of Unstable Multi Output Mechanical Systems," *ASME Paper no. 63-WA-148*, 1963.

J-1 Jury, E. I., *Sampled-Data Control Systems*, New York, N.Y.: John Wi & Sons, Inc., 1958.

J-2 Jury, E. I., "A Simplified Stability Criterion for Linear Discrete System *Proc. IRE*, **50** (1962), pp. 1493–500.

J-3 Jury, E. I., *Theory and Applications of z-Transform Method*, New Yor N.Y.: John Wiley & Sons, Inc., 1964.

K-1 Kalman, R. E., "Analysis and Design Principles of Second and High Order Saturating Servomechanisms," *AIEE Trans. Part II*, **74** (1955), p 294–310.

K-2 Kalman, R. E., "Contributions to the Theory of Optimal Control," *Bo Soc. Mat. Mex.*, **5** (1960), pp. 102–19.

K-3 Kalman, R. E., "On the General Theory of Control Systems," *Proc. Fir Intern, Cong. IFAC*, Moscow, 1960; *Automatic and Remote Control*, Londo England: Butterworths & Company, Ltd., 1961, pp. 481–92.

K-4 Kalman, R. E., "Canonical Structure of Linear Dynamical Systems," *Pro Natl. Acad. Sci., USA* **48** (1962), pp. 596–600.

K-5 Kalman, R. E., "When Is a Linear Control System Optimal?" *ASME J Basic Engineering*, ser. D, **86** (1964), pp. 51–60.

K-6 Kalman, R. E., and J. E. Bertram, "Control System Analysis and Desig Via the Second Method of Lyapunov: I Continuous-Time Systems," *ASME J. Basic Engineering*, ser. D, **82** (1960), pp. 371–93.

K-7 Kalman, R. E., Y. C. Ho, and K. S. Narendra, "Controllability of Linea Dynamic Systems," in *Contributions to Differential Equations*, vol. 1, New York, N.Y.: Interscience Publishers, Inc., 1962.

K-8 Korn, G. A., and T. M. Korn, *Electronic Analog Computers*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1956.

K-9 Kreindler, E., "Contributions to the Theory of Time Optimal Control," *J. Franklin Inst.*, **275** (1963), pp. 314–44.

K-10 Kreindler, E., and P. E. Sarachick, "On the Concepts of Controllability and Observability of Linear Systems," *IEEE Trans. Automatic Contro* (1964), pp. 129–36.

K-11 Kuo, B. C., *Automatic Control Systems*, Englewood Cliffs, N.J.: Prentice Hall, Inc., 1962.

K-12 Kuo, B. C., *Analysis and Synthesis of Sampled-Data Control Systems*, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963.

L-1 Lago, G., and L. M. Benningfield, *Control System Theory*, New York, N.Y.: Ronald Press Company, 1962.

L-2 LaSalle, J. P., "Time Optimal Control Systems," *Proc. Natl. Acad. Sci., USA*, **45** (1959), pp. 573–7.

L-3 LaSalle, J. P., "Time Optimal Control," *Bòl. Soc. Mat. Mex.*, **5** (1960), pp. 120–4.

L-4 LaSalle, J. P., "The Time-optimal Control Problem," in *Contributions to Differential Equations*, vol. 5, Princeton, N.J.: Princeton University Press, 1960, pp. 1–24.

L-5 LaSalle, J. P., "The Bang-Bang Principle," *Proc. First Intern. Cong. IFAC*, Moscow, 1960; *Automatic and Remote Control*, London, England: Butterworths & Company, Ltd., 1961, pp. 493–7.

L-6 LaSalle, J. P., and S. Lefschetz, *Stability by Liapunov's Direct Method with Applications*, New York, N.Y.: Academic Press, Inc., 1961.

L-7 Lefschetz, S., *Differential Equations: Geometric Theory*, New York, N.Y.: Interscience Publishers, Inc., 1957.

L-8 Lewis, J. B., "The Use of Nonlinear Feedback to Improve the Transient Response of a Servomechanism," *AIEE Trans. Part II*, **71** (1952), pp. 449–53.

L-9 Lindorff, D. P., *Theory of Sampled Data Control Systems*, New York, N.Y.: John Wiley & Sons, Inc., 1965.

M-1 Mason, S. J., "Feedback Theory: Some Properties of Signal Flow Graphs," *Proc. IRE*, **41** (1953), pp. 1144–56.

M-2 Mason, S. J., "Feedback Theory: Further Properties of Signal Flow Graphs," *Proc. IRE*, **44** (1956), pp. 920–6.

M-3 Mason, S. J., and H. J. Zimmerman, *Electronic Circuits, Signals, and Systems*, New York, N.Y.: John Wiley & Sons, Inc., 1960.

M-4 Melbourne, W. G., "Three Dimensional Optimum Thrust Trajectories for Power-Limited Propulsion Systems," *ARS J.*, **31** (1961), pp. 1723–8.

M-5 Melbourne, W. G., and C. G. Sauer, Jr., "Optimum Interplanetary Rendezvous with Power-Limited Vehicles, *AIAA J.*, **1** (1963), pp. 54–60.

M-6 Minorsky, N., *Nonlinear Oscillations*, Princeton, N.J.: D. Van Nostrand Company, Inc., 1962.

M-7 Mirsky, L., *An Introduction to Linear Algebra*, Oxford, England: Clarendon Press, Ltd., 1955.

M-8 Monopoli, R. V., "Controller Design for Nonlinear and Time-Varying Plants," *NASA CR*-**152**, Jan., 1965.

N-1 Nyquist, H., "Regeneration Theory," *Bell System Tech. J.*, **11** (1932), pp. 126–47.

O-1 Ogata, K., *State Space Analysis of Control Systems*, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1967.

P-1 Peschon, J., "Learning in Automatic Systems," *Proc. Allerton Conference on Circuit and System Theory*, Univ. of Illinois, Sept., 1964.

P-2 Pipes, L. A., *Matrix Methods for Engineers*, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963.

R-1 Ragazzini, J. R., and G. F. Franklin, *Sampled Data Control Systems*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1958.

R-2 Reid, W. T., "A Matrix Differential Equation of Riccati Type," *Amer. J. Math.*, **68** (1946), pp. 237–46.

R-3 Rekasius, Z. V., "A General Performance Index for Analytical Design of Control Systems," *IRE Trans. Automatic Control* (1961), pp. 217–22.

S-1 Schultz, D. G., and J. E. Gibson, "The Variable Gradient Method for Generating Liapunov Functions;" *AIEE Trans. Part II*, **81** (1962), pp. 203–9.

S-2 Schultz, W. C., and V. C. Rideout, "Control System Performance Measures: Past, Present, and Future," *IRE Trans., Automatic Control* (1961), pp. 22–35.

S-3 Shannon, C. E., "Communication in the Presence of Noise," *Proc. IRE*, **37** (1949), pp. 10–21.

S-4 Smith, O. J. M., *Feedback Control Systems*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1958.

S-5 Staats, P. F., "A Survey of Adaptive Control Topics'" *Plan B paper*, Dept. of Mech. Eng., Univ. of Minn., Mar., 1966.

S-6 Stallard, D. V., "A Series Method of Calculating Control System Transient Responses from the Frequency Response," *AIEE Trans., Part II*, **74** (1955), pp. 61–4.

T-1 Thaler, G. J., and R. G. Brown, *Analysis and Design of Feedback Control Systems*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1960.

T-2 Thaler, G. J., and M. P. Pastel, *Nonlinear Control Systems*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1962.

T-3 Tou, J. T., *Digital and Sampled-data Control Systems*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1959.

T-4 Tou, J. T., *Modern Control Theory*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1964.

T-5 Tsien, H. S., *Engineering Cybernetics*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1958.

T-6 Turnbull, H. W., and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices*, London, England: Blackie and Son, Ltd., 1932.

V-1 Van Landingham, H. F., and W. A. Blackwell, "Controller Design for Nonlinear and Time-Varying Plants," *Educational Monograph*, College of Engineering, Oklahoma State Univ., 1967.

W-1 Wadel, L. B., "Describing Function as Power Series," *IRE Trans. Automatic Control* (1962), p. 50.

W-2 Waltz, M. D., and K. S. Fu, "A Learning Control System," *Proc. Joint Automatic Control Conference*, 1964, pp. 1–5.

W-3 Wilcox, R. B., "Analysis and Synthesis of Dynamic Performance of Industrial Organizations—The Application of Feedback Control Techniques to Organizational Systems," *IRE Trans. Automatic Control* (1962), pp. 55–67.

W-4 Wojcik, C. K., "Analytical Representation of the Root Locus," *ASME J. Basic Engineering*, ser. D, **86** (1964), pp. 37–43.

Z-1 Zadeh, L. A., and C. A. Desoer, *Linear System Theory; The State Space Approach*, New York, N.Y.: McGraw-Hill Book Company, Inc., 1963.

Índice Alfabético

Os números em *itálico* referem-se a localizações fora do texto (legendas, quadros, dísticos, notas, etc.).

A

B

C

H

I

K

L

M

N

O

P

Q

R

S